Title: Characterisation of the novel TPI-GAPDH fusion enzyme in Blastocystis hominis

Name: Kayleigh Evans

CHARACTERISATION OF THE NOVEL TPI-GAPDH FUSION ENZYME IN *BLASTOCYSTIS HOMINIS*

by

Kayleigh Evans

A thesis submitted to the University of Bedfordshire, in partial fulfilment of the requirements for the degree of Master of Science by research.

March 2014

Abstract

Protist parasites such as the human pathogens *Trypanosoma brucei, Blastocystis hominis* and the plant pathogen *Phytophthora infestans* cause devastating diseases worldwide and in the case of *P. infestans*, huge financial losses in potato crops. Treating these diseases is fraught with toxic side effects due to the similar functioning of all eukaryotes and emerging resistance problems. This highlights a requirement for new and innovative antiparasitic treatments that are specific against the target organism without damaging the host. Studies haves demonstrated protist metabolism as a potential drug target.

A novel mitochondrial targeted glycolytically active TPI-GAPDH fusion enzyme has been discovered in a selection of stramenopiles (including *B. hominis* and *P. infestans*). This is interesting in terms of a potential selective drug target and when considering the mitochondrial targeting to the enzyme, suggesting a retention of glycolysis originating from the mitochondrial precursor before the endosymbiosis event (i.e. from *alpha proteobacteria*).

This study characterises TPI-GAPDH from *B. hominis* by performing enzymatic assays on wild type TPI-GAPDH, mutated TPI-GAPDH and by splitting the protein into separate TPI and GAPDH enzymes. The enzymatic assays showed that GAPDH does not function as effectively when separated from TPI, having 2-fold decrease in enzyme activity compared against wtTPI-GAPDH. This suggests that TPI has a structural effect on GAPDH function. The enzymatic assays also demonstrated a competition for GAP between TPI and GAPDH, this goes against the canonical understanding of TPI. TPI functions to equilibrate DHAP to GAP in a 22:1 ratio and is generally considered to be a 'perfect' enzyme which is not rate limited in its function. However, when attached to GAPDH, TPI appears to be in competition with GAPDH for substrate. The apparent loss of function of *B. hominis* GAPDH when separated from TPI makes the linkage between the two enzymes a potential novel drug target within glycolysis in these protists.

Author's Declaration

I declare that this thesis is my own unaided work. It is being submitted for the degree of Master of Science by research at the University of Bedfordshire.

It has not been submitted before for any degree or examination in any other University.

Name of candidate: Kayleigh Evans Signature:

Date: 10/03/14

Dedication

Dedicated to my friends and family, especially Michael, Emma and Viki. Thank you for all your love, support and patience. I wouldn't be me without you.

And Robin, who let me play in his lab and encouraged me to do this. Thank you, it was worth it.

Table of Contents

List of Figures and Tables

Acknowledgments

I would like to thank:

- My supervisor Dr Robin Maytum for all his support, guidance, encouragement, teachings and hot chocolate throughout the year.
- Michael Evans for being an Excel god and supporting me throughout all of my studies
- Emma Amir-Embrahimi and Viki Boncheva for keeping me sane and keeping me laughing
- Khadar Dudekula, for all her help and support in the lab
- Dr Mark Van Der Giezen of University of Exeter for providing us with the *B. hominis* TPI-GAPDH, TPI and GAPDH clones

Abbreviations

1,3 BPG - 1,3-Bisphosphoglyceric acid

Amp - Ampicillin

B. hominis – Blastocystis hominis

[E] – Enzyme concentration

GAPDH - glyceraldehyde-3-phosphate dehydrogenase

HAT – Human African Trypanosomiasis

IPTG - Isopropyl β-D-1-thiogalactopyranoside

LB – Lysogeny Broth

MWT – Molecular weight

OD – Optical Density

P. infestans – Phytophthora infestans

SDS-PAGE - sodium dodecyl sulfate polyacrylamide gel electrophoresis

[S] – Substrate concentration

T. brucei – Trypanosoma brucei

TPI - triose phosphate isomerase

TPI-GAPDH – Triose phosphate isomerase glyceraldehyde-3-phosphate dehydrogenase fusion protein

T. vaginalis – Trichomonas vaginalis

wtTPI-GAPDH – Wild type triose phosphate isomerase glyceraldehyde-3-phosphate dehydrogenase fusion protein

Chapter 1: Introduction

1.1 Protist Parasites

1.1.1 Protist Classification and Physiology

Protists are a diverse group of unicellular eukaryotic organisms. They are often only distantly related to each other however have historically been placed in the same kingdom as they do not fit morphologically elsewhere. The only true similarity between protists is that they are unicellular and eukaryotic. However, some protists exhibit high levels of similarity to other eukaryotes, so their classification is often controversial, for example the Stramenopile *Phytophthora infestans* has historically been classified as a fungi, due to it's highly similar morphology and life cycle to fungi. However, phylogenetic analysis shows *P. infestans* to be genetically closer to other protist organisms and diverged away from fungi very early on in its evolution (Haas et al., 2009). Modern classification only uses the term protists as a convenience to refer to the unicellular eukaryotes; protists are now vaguely separated into super-groups as separate lines of the eukaryote branch rather than an entire kingdom (Wiser, 2011). For convenience, this paper will still refer to unicellular eukaryotes as protists.

The phylogenetic classification of protists has been a difficult task and still continues. Phylogenetic trees for protists are constantly being changed and updated as the range of genes sequenced for phylogenetic analysis increases. This is partly due to the fact that phylogenetic analysis is generally done on the 18S small subunit rRNA genome. Some protists appear to have undergone very early evolutionary branching, periods of rapid evolution and have the ability to acquire

genes by lateral gene transfer (LGT). This divergence and evolution have led to artefacts from other organisms appearing in areas that are meant to be highly conserved, confusing the mathematical models of phylogenetic analysis. Phylogenetic analysis is now done on a range of highly conserved genes, such as, tubulin, RNA polymerase and ATPase in the hopes of a more accurate picture of the evolution of an organism. Whilst phylogenetic analysis is useful, there are limitations; there are still not many full genomes for protists available for sequencing and analysis. It is also becoming more apparent that the mathematical models used to simulate phylogenetic trees are too simplistic to emulate sequence evolution, i.e. different sequence sites in genomes evolve at different rates; leading to simulated models that produce a phylogenetic tree with the wrong links (such as long branch attraction – fast evolving organisms in relation to others in the tree become clustered together irrespective their other similarities or differences) but which have strong modelling evidence to support them (Embley and Martin, 2006; Roger and Hug, 2006; Tamura et al., 2012).

The diversity of protists becomes evident on closer inspection of the organisms, whilst they retain certain cellular level similarities (e.g. double membrane bound organelles) with other eukaryotes, there are also plenty of differences. Many protists show diversity in their organelles; some have dimorphic nuclei, some have mitochondria or plastids, whilst others have lost theirs or retain part functioning mitochondrial/plastid remnants, others have unique organelles such as apicoplasts (Plasmodium) and glycosomes (Trypanosoma) (Ginger, 2006). This allows protists to be relatively complex. The range of structures available to them is indicative of their lifestyle as single celled organisms; they do not have the luxury of delegating tasks to differentiated cells. Unlike multicellular eukaryotes they must do everything in the one cell or depend on their host for certain functions, such as nutrient acquisition. The variation in protists is also due to them being environmentally ubiquitous; the universality of the organisms has led to extensive and peculiar adaptions to environmental niches – from metabolism to life cycles with very specific stages designed for their current habitat (Müller et al., 2012). Whilst most protists tend to prefer aquatic environments, there is a

significant number that are pathogenic parasites of plants and animals, including humans, of which are the cause of some of the most prevalent human and plant diseases.

1.1.2 Treating Protist Parasite Diseases

Treating bacterial infection, whilst fraught with resistance problems and patients not taking medication properly, is comparatively easy next to treating protist infections. Bacterial infections can often be treated with antibiotics; the range of which covers inhibition of: cell wall synthesis (e.g. penicillin and vancomycin), protein synthesis (e.g. tetracyclins and chloramphenicol) and nucleic acid synthesis (e.g. quinolones). This arsenal of drug therapies against bacterial disease allows for the majority of infections to be cured. The differences being that the mechanisms of antibiotic action are targeted specifically to structures within prokaryotic cells (Madigan et al., 2009). Protist parasites are eukaryotic cells, therefore drug therapy involving, for example, disrupting the cellular membrane would likely result in disrupting the cellular membrane of the host cells, leading to the toxic side effects that anti-parasitic drugs are known for.

The toxicity of drug therapies against protist parasites is a major problem in combatting diseases caused by them. One notable example is the treatment of Human African Trypanosomiasis (HAT), caused by *Trypanosoma brucei* (*T. brucei)* (Figure 1). A fatal disease found in sub-Sahara Africa transmitted by an insect vector, the Tsetse fly. Around 60 million people are believed to be “at risk” of this disease; which causes headaches, malaise and fever in the early stages then develops into progressively worsening neurological symptoms, involving all regions of the nervous system and eventually disseminating into multiple organ involvement, including significant cardiac problems.

Figure 1: Bloodstream form of *T. brucei* from a blood smear of a patient with HAT.
From M. Schultz. Centres for Disease Control Public Health Image Library.

The treatment of HAT is not ideal; it relies on drugs that are decades old (Suramin has been in use since 1920 and Melarsoprol since 1949), with unknown mechanisms of action that are highly toxic to the patients. Suramin and Pentamidine are used to treat the early stage HAT and though they are usually effective they offer the patient a range of potential side effects including renal failure, anaphylactic shock, neurological complications and hypotension. Treating late stage HAT is much more dangerous to the patient, as Melarsoprol is the only treatment available to those with *T. brucei rhodesiense* and is the 2nd line of treatment in *T. brucei gambiense* (there are 2 sub-species of *T. brucei*, both affecting different regions of sub-Saharan Africa with *T. brucei rhodesiense* being the more aggressive species). Melarsoprol is an arsenical compound that is extremely toxic, so much so 5.9% of patients die from complications (mainly reactive encephalopathy); other side effects include neuropathy, multifocal inflammatory disorder, hypertension and significant heart damage (Jacobs et al., 2011; Kennedy, 2013). The potential of a drug therapy to kill the patient is a clear sign that better treatment needs to be found.

Treatment of HAT is not the only chemotherapy against protists to have side effects. Whilst not as severe as the treatment for HAT, the side effects from the treatment of a range of other protist caused diseases are still dangerous. Benznidazole, used to treat acute Chagas' Disease (caused by *Trypanosoma cruzi*)

can cause polyneuropathy, polyneuritis and bone marrow disorders, possibly due to the mechanism of action of the drug (reductive stress mechanism) causing oxidative/reductive stress on the host tissue. There is no recommended anti-parasitic drug for chronic Chagas' Disease (Salomon, 2012).

Whilst the difficulty in treating these two major neglected diseases is significant, there are many challenges still to face with some of the most common diseases all over the world. *Blastocystis hominis (B. hominis)* (Figure 2) and *Giardia lamblia* are two of the most frequently isolated pathogens from stools samples when investigating the cause of diarrhoea and other gastrointestinal problems, with *B. hominis* being the more common (Omar, 2006; Tan, 2008).

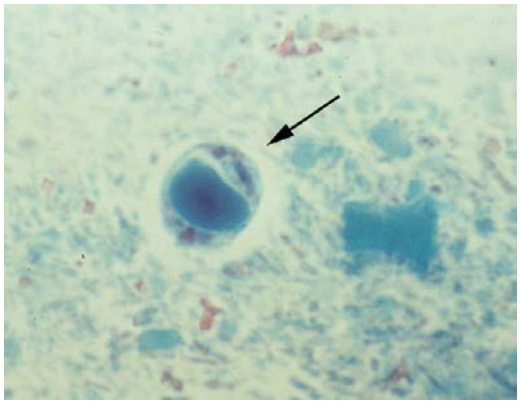


Figure 2: Trichrome stain showing a Blastocystis cyst in stool.
From the H. Zaiman. American Society of Tropical Medicine and Hygiene.

The prevalence of *B. hominis* varies dramatically both between countries and within countries, with less developed areas having a higher incidence. For example, a study in Japan only identified 0.5% of the population infected with *B. hominis* (Hirata et al., 2007), whilst Pegelow et al. found the parasite present in 60% of schoolchildren in Indonesia (Pegelow et al., 1997). Variability across countries can range dramatically also, such as in China, where a study found the incidence ranging between 1.9% in Shanghai compared to 32.6% found in Menghai county, reflecting the contrast between the development of different areas of the country (Li et al., 2007). Coupling the worldwide distribution of *B. hominis* with the

fact that the protist can be both symptomatic and asymptomatic makes controlling the disease a significant challenge. Treatment for *B. hominis* is a controversial issue as there is no real consensus on whether the microorganism is pathogenic or not considering the large variety of gut microbiota (Tan, 2008; Yakoob et al., 2004). Although a recent paper does suggest that *B. hominis* proteases do exacerbate the symptoms caused by a gastrointestinal infection, suggesting that *B. hominis* can be the causative agent in some cases (Rajamanikam and Govind, 2013). Regardless, there is a response in some patients to the broad-spectrum antibiotic metronidazole, which targets anaerobic organisms in the gut. However this is not always effective and resistance to metronidazole is being reported (Haresh et al., 1999; Yakoob et al., 2004). It is thought that metronidazole works by inhibiting DNA synthesis and causing damage to DNA by oxidation, however the mechanism of action is not fully understood yet (Löfmark et al., 2010). Whilst most patients suffer few side effects from the drug if any, there have been some reported cases of serious neurological side effects such as seizures and encephalopathy that cease once treatment stops (Sarna et al., 2013). Like treatment for HAT (though not as serious), the ambiguity of the mechanism of drug action and the potential side effects, coupled with the general difficulty in treating *B. hominis* highlights the need not only for a more effective form of treatment but also a more specific treatment to cause minimum damage to the patient.

The trend set by *B. hominis* is continued by many of the other human pathogenic protists; *Giardia lamblia* is an equally common cause of gastrointestinal disease but equally as difficult to treat, as well as developing resistance to metronidazole (Lalle, 2010). *Trichomonas vaginalis*, is one of the most common STIs worldwide, due to often being asymptomatic and going unnoticed. It can lead to a serious increased risk of cervical cancer and HIV susceptibility, as well as premature labour in pregnant women (Cudmore et al., 2004). It is also difficult to diagnose, eradicate effectively and is gaining resistance to metronidazole. Again pointing towards the need for more effective antiparasitic drugs.

Protist parasites are also very significant plant pathogens. The most notorious being *Phytophthora infestans* (Figure 3), the causative agent of the Irish Potato Famine of the 1840s that led to the deaths of over 100 million Irish by starvation. *P. infestans* continues to blight potato and tomato crops today causing an estimated cost of $6.7 billion in damages and control mechanisms (Haverkort et al., 2008; Strange and Scott, 2005). The species is morphologically fungus-like in its lifecycle and in its ability to produce spores that can be easily dispersed through wind and water, as well as survive in unfavourable conditions. Whilst there are a lot of similarities to fungi, *P. infestans* and other Oomcyetes branched away from plants and fungi very early on and are now regarded to belong to the Stramenopiles group, like *B. hominis*. This early divergence means *P. infestans* has evolved quite differently from plants and fungi, ensuring that most plant defences against fungi are ineffective against the protist, for example, chitinases will break down the cell wall of fungi but Oomcyetes contain no chitin in their cells (Fry, 2008).

Figure 3: Sporangia of *P. infestans* that would be released into the atmosphere to be dispersed by air or water.
From W. Fry. *Phytophthora infestans*: the plant (and R gene) destroyer

Some resistance to *P. infestans* has been achieved in the wild and selective breeding is done to ensure the trait is passed on. However the pathogen has high levels of genetic and phenotypic diversity; it can often come up with new virulence factors quickly. This is largely due to its extensive and fast evolving repertoire of effector genes. These effector genes help in colonisation and destruction of the plant; producing enzymes such as hydrolases (to break down plant matter) and

proteins that modulate plant immunity (Haas et al., 2009). The use of pesticides has been employed help control *P. infestans* on both current crops and those already infected but with the potential of the organism going through multiple infection cycles in a week the benefit of using pesticides is limited (Cooke et al., 2012). The difficulty in controlling this pathogen is evident; it is a significant threat to global food security, with the potential to cause famine to those that depend on the crop in the developing world, more solutions are needed to combat this parasite.

1.1.3 Potential Drug Targets

The limited treatment options for protist parasites due to: the resistance to broad-spectrum drugs, limited vaccine development and toxic side effects, has meant research has had to delve away from looking purely at traditional ways to inhibit proliferation (i.e. inhibition of cell wall synthesis, protein synthesis and DNA synthesis). The ever-increasing wealth of genetic information being gathered on organisms has meant that scientists can look at challenging pathogens on a more specific level (http://www.genedb.org/ contains a range of microbial genomes). Genomes can be searched for potential pathogen specific drug targets (by genome mining) in the hopes to minimise any potential toxicity problems against the host. Therefore targeting something that is unique yet essential to a protist, such as an organelle or enzyme pathway, is preferred. A plethora of potential targets have been found this way, including; targeting gene expression (Bell et al., 1991; Darkin-Rattray et al., 1996; Leepin et al., 2008), the cytoskeleton (Müller and Hemphill, 2011), intracellular signalling (Hammarton et al., 2003; Kunz et al., 2006), membrane integrity (Rayan et al., 2005; Winter et al., 2008), toxic radical formation (Li and Xin, 2009) and metabolism (Cáceres et al., 2010; Jacobs et al., 2011; Mallo et al., 2013; Tielens et al., 2010), to name a selection. Many of these studies have found a unique pathway to exploit, however there has been some discoveries into a more ubiquitous enzyme, glyceraldehyde-3-phosphate dehydrogenase (GAPDH), with a property exclusive to a selection of protists, some of which include pathogenic protists (Liaud et al., 2000; Nakayama et al., 2012).

1.1.4 Protist Metabolism

Eukaryotic organisms share much of the same basic metabolism, with a few interesting exceptions such as obligate intracellular organisms like microsporidia, which have a highly reduced metabolism as they rely on taking nutrients and macromolecular precursors from their host. Eukaryotic metabolism is believed to stem from one common anaerobic ancestor with the entire future metabolic repertoire being a product from that initial ancestor and any evolutionary changes since. Eukaryotic energy metabolism all revolves around some form of carbon metabolism; relatively little has changed in the basic biochemistry and it has either been added by lateral gene transfer or slimmed down by gene loss to fit a niche. This is difference to prokaryotes, which fit almost any environmental niche, therefore are able to metabolise a wider variety of food sources that eukaryotes cannot. Their metabolic ancestry is not as linked as eukaryotic metabolism, further pointing towards a single archeal origin for eukaryotes (Der Giezen, 2011; Embley and Martin, 2006; Ginger, 2006; Müller et al., 2012; Tielens et al., 2010).

The commonality of eukaryotic metabolism does not suggest it would be a good potential drug target, however small differences do provide possibilities. Most drugs do target metabolism, therefore understanding pathogenic organisms' metabolism is always important, especially if they do something unusual. For example, many studies are looking into targeting glycolysis in trypanosomes. Glycolysis is their only source of energy whilst they are in the bloodstream of their host but also because glycolysis occurs in a peroxisome-like glycosome which is unique to the trypanosomes, making this a highly desired target (Bero et al., 2013; Bringaud et al., 2006; Cáceres et al., 2010; Jacobs et al., 2011; Kuntz et al., 1992; Michels et al., 2006). Metabolic organelles that can be found in a range of protists but not multicellular eukaryotes also have the potential to be drug targets. Hydrogenosomes are reduced forms of mitochondria that produce ATP by fermentation and substrate level phosphorylation, they are found among a range

of protists, including *T. vaginalis*. Research into targeting hydrogenosomes has discovered that resveratrol causes hydrogenosomal disruptions that significantly inhibited growth and was cytotoxic to the cells, showing potential for further investigation as an antiparasitic drug and for hydrogenosomes as a drug target (Mallo et al., 2013). Targeting large organelles is not the only research being done; much is looking at single or multiple enzymes involved in metabolism. A surprising target, as mentioned earlier, is GAPDH. Whilst there are studies investigating the enzyme on its own (Bero et al., 2013; Cáceres et al., 2010), recent developments in understanding the evolutionary origin of eukaryotes has unearthed an interesting trait of GAPDH in a selection of protists; it is a fusion protein with triosephosphate isomerase (TPI) and it is mitochondrial targeted. Whilst the evolutionary side is interesting and will be expanded on briefly later, this unique linkage raises some interesting questions. Is this fusion enzyme functional? Does it compare to other GAPDH enzymes from other organisms? How important is it that the enzyme remains a fusion? Is this linkage a drug potential target?

1.2 GAPDH Structure and Function

1.2.1 Glycolytic Function of GAPDH

Before exploring the fusion enzyme, the functioning of GAPDH will first be expanded upon to offer context. GAPDH is one of the most studied enzymes and is used frequently as a housekeeping gene and a baseline enzyme for protein assays. It is a tetramer of 4 identical 37kDa subunits (Figure 4a). Each subunit contains a single catalytic site separated into two binding domains for two molecules: the substrate Glyceraldehyde 3 Phosphate (GAP) and the co-enzyme Nicotinamide adenine dinucleotide (NAD^+) (Figure 4b). The two binding domains form a catalytic site in which NAD^+, a cysteine residue at 149 (C149) and histidine residue at 176 (H176) play key roles in substrate binding (Moras et al., 1975). The active site residues referred to above are those denoted in the archetypal lobster primary sequence as shown in Appendix 1 (Davidson et al., 1967). There can be variation in

primary sequence between organisms; this can be seen in Figure 4c, which shows the active residues of a methicillin resistant *Staphylococcus aureus* to be C151 (instead of C149) and H178 (instead of H176).

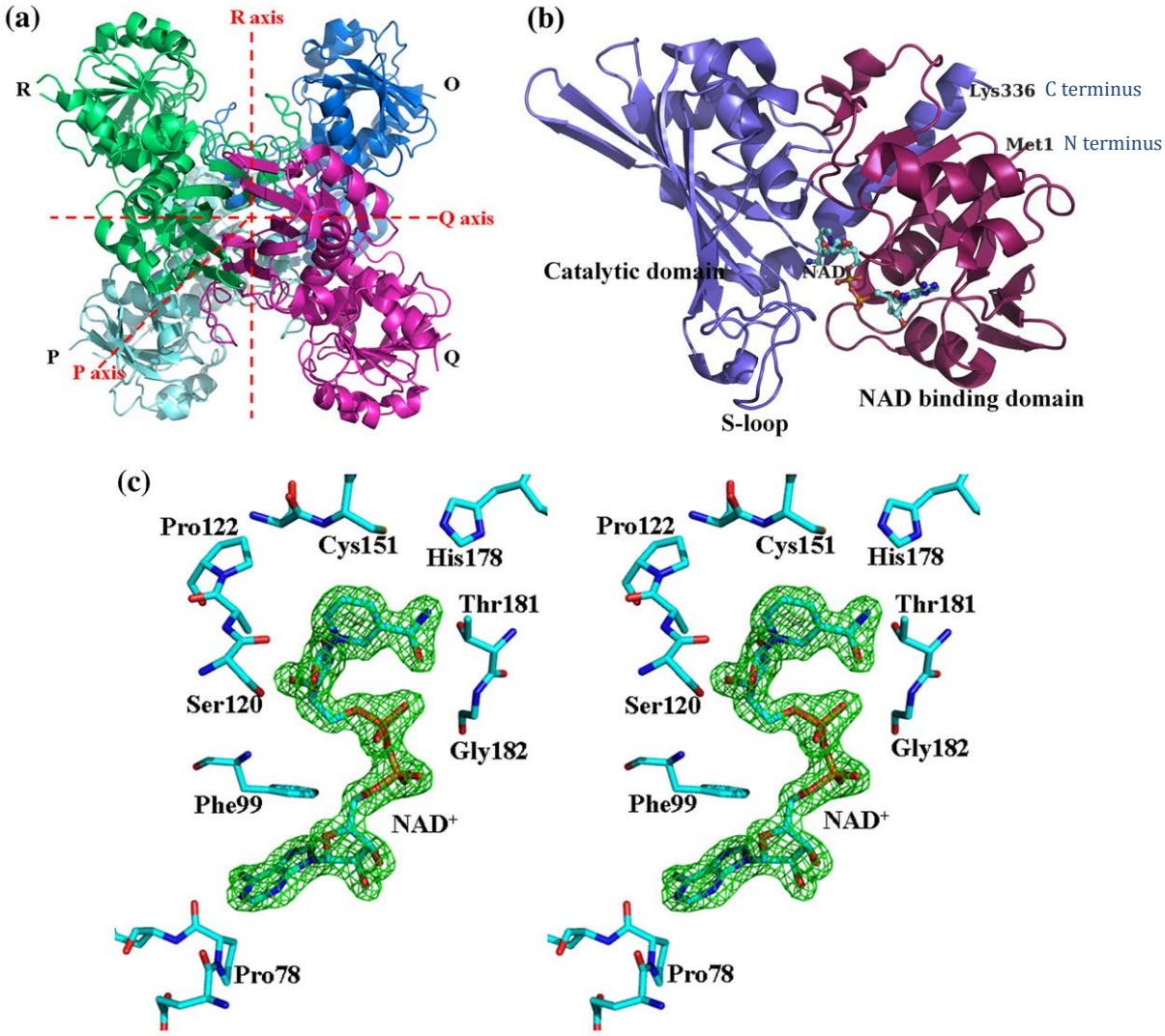


Figure 4: (a) Crystal Structure of tetrameric GAPDH from MRSA. (b) The catalytic domain (purple) binds the substrate and forms the catalytic site with the NAD binding domain (pink) The N terminus at the methionine and the C terminus at the lysine are also highlighted. (c) Cys151 and His178, along with NAD+ form the substrate binding site of MRSA GAPDH. Other interacting residues are present.

From Mukherjee *et al*. 2010. Crystal Structure of Glyceraldehyde-3-Phosphate Dehydrogenase 1 from Methicillin-Resistant Staphylococcus aureus MRSA252 Provides Novel Insights into Substrate Binding and Catalytic Mechanism

The classic function of GAPDH is as a glycolytic enzyme in the cytosol of cells; it has been well characterised in this role since the 1950s (Oesper, 1954; Trentham, 1968). Glycolysis is the first stage of energy metabolism from glucose, producing 2 pyruvate molecules that are then fed into the Kreb's Cycle (Figure 5).

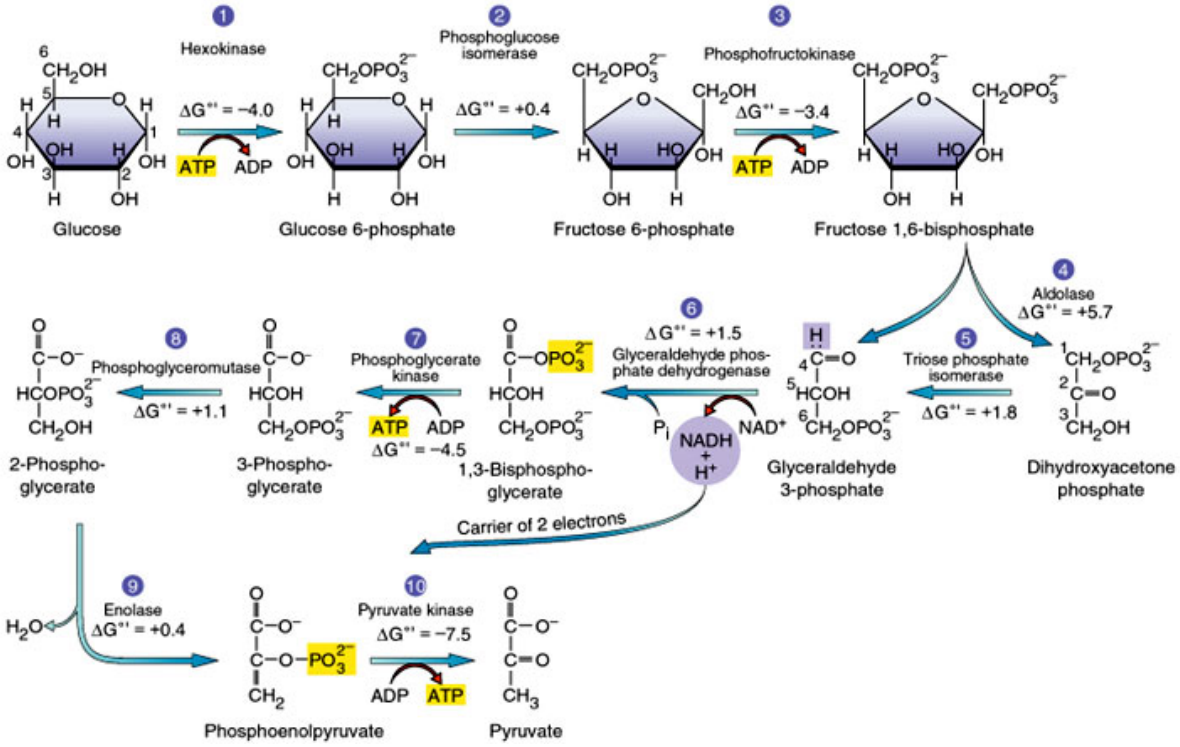


Figure 5: Glycolysis

From Karp, G. (2009). Cell and Molecular Biology: Concepts and Experiments, 5th ed.

Glucose enters the cell and is trapped in the cytosol when converted into glucose 6-phosphate by hexokinase (step 1); the phosphorylation makes the glucose hydrophilic so it cannot cross back through the cell membrane. From there it continues through a series of preparation steps until the reaction with aldolase (step 4), which splits the fructose 1,6-bisphosphate 50:50 into glyceraldehyde 3-phosphate (GAP) and dihydroxyacetone phosphate (DHAP). Triose phosphate isomerase (TPI) (step 5) converts between the two, as only GAP can continue along the pathway (more will be explained on the dynamics of TPI later). TPI

equilibrates DHAP and GAP at their thermodynamic equilibrium of 22:1 within the cell, thus limiting the rate of GAPDH. GAP is converted to 1,3 bisphosphoglycerate (1, 3 BPG) by the oxidoreduction reactions preformed by GAPDH (step 6). This is a vitally important step of glycolysis; it produces the product for the next step of glycolysis, which is the first ATP producing step by substrate level phosphorylation (if both GAPs from one glucose molecule are taken into account, then the two ATPs invested in steps 1 and 3 are repaid). It also produces two NADH + H^+ from NAD^+ for each molecule of glucose. In aerobic respiration, the NADH is transported to the mitochondria where it takes part in further ATP synthesis by oxidative phosphorylation and is oxidised back to NAD^+ for use again in glycolysis and the Krebs cycle (aerobic respiration). However, for those organisms that are amitochondriate, have reduced mitochondrial remnants or go through periods of oxygen starvation; NAD^+ is regained through fermentation reactions that result in the oxidation NADH + H^+ into NAD^+. From phosphoglycerate kinase (step 7), glycolysis continues for 3 more steps until ultimately two molecules of pyruvate and two further ATP molecules are produced from a single glucose. Producing a net gain of 2 ATP molecules for each molecule of glucose (Berg et al., 2007; Harris, 2006; Mukherjee et al., 2010).

The reaction catalysed by GAPDH is an oxidoreduction reaction, the oxidation reaction being favourable and allowing the reduction reaction to happen. The mechanism of the enzyme begins with C149 (Cys) being deprotonated by the basic H176 (His) to form a thiolate anion (S^-). The thiolate performs a nucleophilic attack on the aldehyde group of GAP. This forms a thiohemiacetal intermediate between the substrate and sulfhydryl group of the enzyme. The intermediate is oxidised to form a thioester by the transfer of a hydride to NAD^+ to form NADH + H^+. NADH is released from the active site to allow another NAD^+ to bind. The thioester undergoes another nucleophilic attack, this time by inorganic phosphate to form 1,3 BPG and is released from the enzyme. This release makes GAPDH available for another GAP molecule (Figure 6) (Moras et al., 1975; Mukherjee et al., 2010; Oesper, 1954).

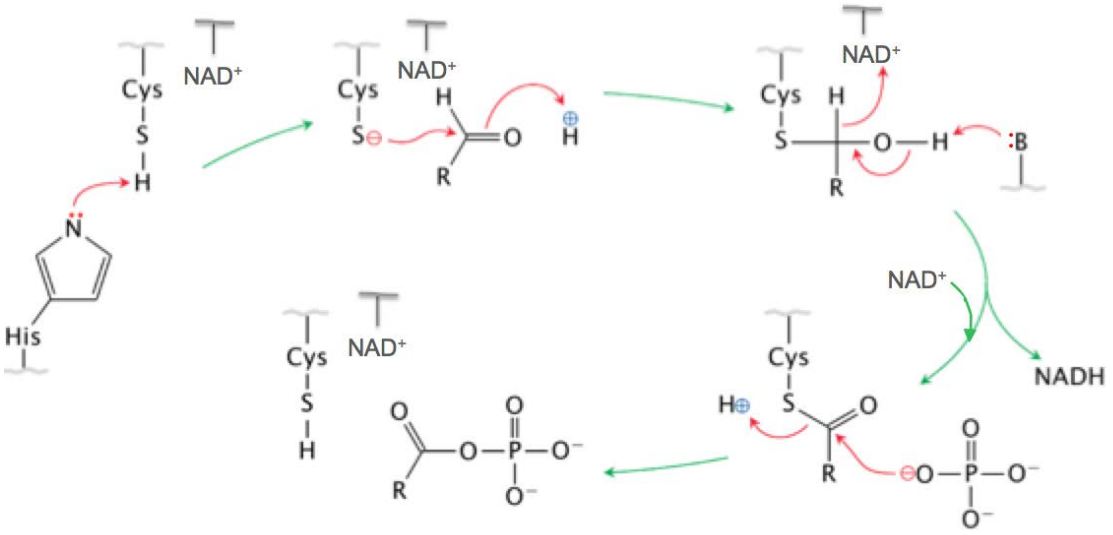


Figure 6: GAPDH mechanism of action on GAP showing the catalytic action of the histidine and cysteine residues, as well as the action of the co-enzyme NAD.

The catalytic residue C149 was targeted in this project, as this is where the enzyme-substrate complex forms during the reaction. As GAPDH is attached to TPI in the wtTPI-GAPDH fusion protein, the position of C149 is shifted along in the primary amino acid sequence to accommodate the TPI moiety. In TPI-GAPDH, GAPDHs active cysteine is at C404 as can be seen in the TPI-GAPDH primary sequence in Appendix 3.

1.2.2 Other GAPDH functions

GAPDH also appears to be a moonlighting protein; recent studies have shown GAPDH to have many other functions outside of glycolysis (Tristan et al., 2011). All other functions appear to be related to maintaining the homeostasis of the cell in some way, such as: modulating cellular signalling pathways during oxidative stress to protect the cell (Kim et al., 2003), promoting apoptosis when oxidative stress becomes too high and damaging (Nakajima et al., 2009; Sirover, 1997), it has been indicated in vesicle trafficking and membrane fusion (Glaser et al., 2002; Tisdale et

al., 2009), expression in the mitochondria has been linked to induction of apoptosis proteins (Tarze et al., 2007), and in the nucleus, glycolytically inactive monomeric GAPDH has been associated with DNA repair (Ronai, 1993). On top of this, GAPDH is thought to play a key role in the neurodegenerative diseases such as Parkinson's and Alzheimer's (Chuang et al., 2005). There have also been reports of simpler protists having multifunctional GAPDH enzymes: *Entaombea histolytica* have been found to apply post translational modifications to their GAPDH, which are possibly involved in signal transduction to host cells and pathogenicity (Alvarez et al., 2007). This is a non-exhaustive list of the extensive range of functions, as well as the primary glycolytic role. This highlights the importance of GAPDH as a potential drug target, if made specific enough from the host.

1.3 TPI Structure and Function

Whilst the focus of this project is on the effects on the activity of the GAPDH enzyme, an understanding of TPI is important, as knocking out the TPI component of TPI-GAPDH allows us to establish the interdependence and functioning of the components of the fusion protein.

TPI is a critical dimeric glycolytic enzyme that catalyses the reversible interconversion between GAP and DHAP. The dimer contains two subunits of 28kDa with the catalytic site at the dimer interface. Two key amino acid residues are involved in the isomerisation reaction: Glutamic acid 167 (E167) and histidine 95 (H95) (Banner et al., 1976). These residues can be seen in the primary amino acid structure of TPI in Appendix 2. E167 is involved in the initiation of the reaction by removal of a proton from the substrate, as well as completing the reaction by transferring the proton back to the isomerised product, whilst H95 is involved in proton mediation (Figure 7). Computer modelling calculates that there are two possible ways in which the reaction happens: the criss-cross mechanism, in which all proton transfer is carried out by E167 and the classical mechanism in which both E167 and H95 are involved (Banerjee et al., 2009; Knowles and Albery,

1977; Samanta et al., 2011; Wierenga et al., 2010). Mutations in both of these residues were investigated. In *B. hominis* TPI-GAPDH, TPIs active residues are at H96 and E167 as can be seen in the TPI-GAPDH primary sequence in Appendix 3.

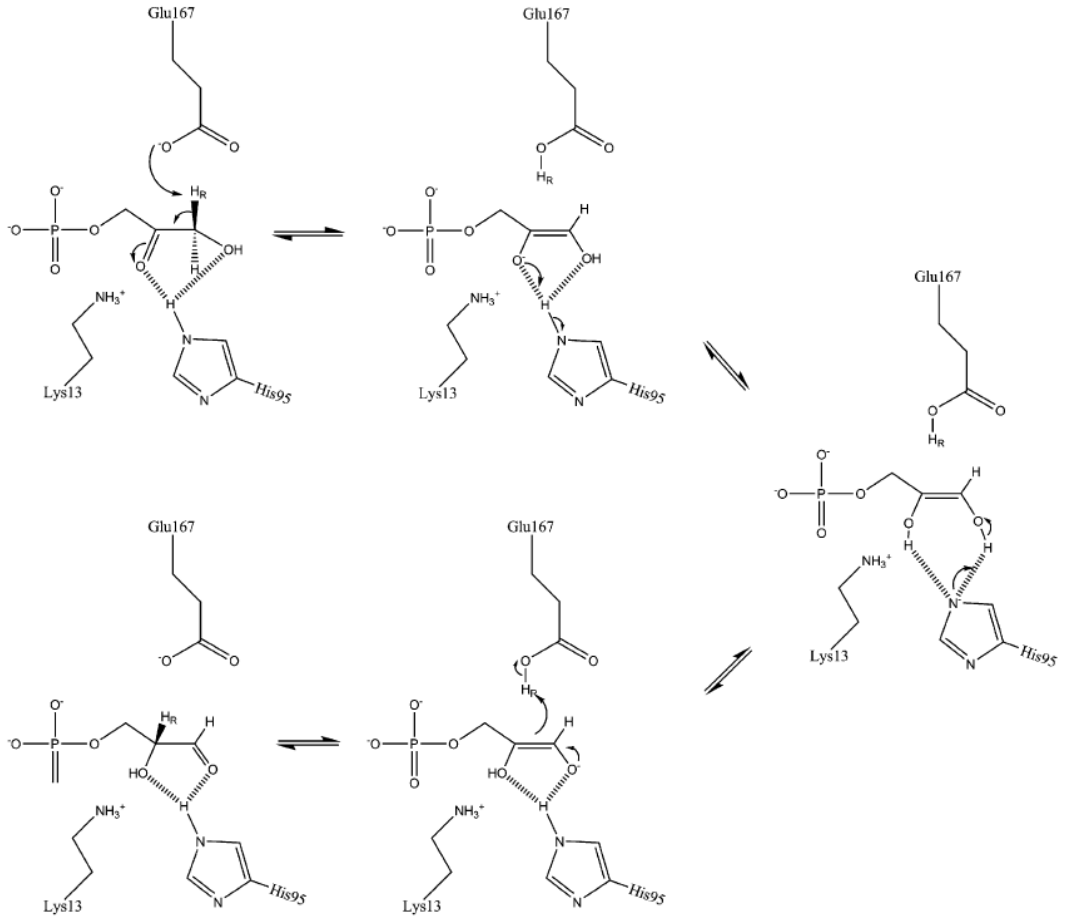


Figure 7: The classical reaction mechanism of TPI showing the proton transfer being carried out by both E167 and H95.
From Wierenga et al. 2010. Triosephosphate isomerase: a highly evolved biocatalyst.

Based in the cytosol of the cell (as well as the glycosome in *Trypanosomes*), TPI is vital to survival; it converts the DHAP from hydrolysis of fructose 1,6-bisphosphate into GAP. However, the thermodynamic equilibrium is 22:1 of DHAP:GAP within the cell, it effectively converts GAP produced from the previous step into DHAP before GAPDH has a chance to utilise it (Knowles and Albery, 1977;Wierenga et al.,

2010). The importance of TPI is highlighted by the fact it is the only human glycolytic enzyme in which deficiency is lethal (Orosz et al., 2009).

1.4 TPI-GAPDH

1.4.1 TPI-GAPDH Structure and Function

The novel TPI-GAPDH fusion protein found in *Stramenopiles*, *Oomcyetes* and other protists has yet to be characterised fully. The structure is still a controversial issue; we believe it to be a tetramer made up of one dimer of TPI and one dimer of GAPDH based on our preliminary size exclusion chromatography data. A study by Liaud *et al* has tentative data suggesting it to be a tetramer of GAPDH with two dimers of TPI either end (Liaud et al., 2000). The latter having more merit as until recently it was believed that all active GAPDH enzymes were tetramers. We now know that there are functional monomers and dimers of GAPDH, although none of these have been shown to have glycolytic activity (Tristan et al., 2011). As TPI has its active site at the dimer interface, it is not thought that it can function in other structural forms (Zhang et al., 1994). Genomic studies have shown TPI and GAPDH to be on the same operon with a full TPI sequence upstream of GAPDH for *B hominis*. TPI-GAPDH is coded so that the C-terminus of the TPI attaches to the N-terminus of the GAPDH. Interestingly, genomic analysis has also uncovered a mitochondrial targeting pre-sequence as part of the fusion protein that has more sequence homology to *alpha-proteobacterial* sequence than to anything in the cytosol. As glycolysis occurs in the cytosol, mitochondrial targeting is strange. It is believed that TPI-GAPDH undertakes the same glycolytic function as the canonical TPI and GAPDH (Liaud et al., 2000; Nakayama et al., 2012).

1.4.2 TPI-GAPDH Evolutionary Implication

The apparent glycolytic function of this fusion enzyme leads to interesting evolutionary ideas into a mitochondrial origin for glycolysis, indeed other glycolytic enzymes (e.g. phosphoglycerate kinase) have been found to be mitochondrial targeted in *Stramenopiles*. The broad distribution in *Stramenopiles* suggests that TPI-GAPDH was present in the ancestral species of modern *Stramenopiles* and potentially goes back to a common ancestor. This possibly suggests a loss of the fusion into two separate enzymes over time for other organisms or alternatively a loss of the *alpha-proteobacterial* form (i.e. mitochondrial form). Phylogenetic analysis on TPI-GAPDH in *Stramenopiles* found both the TPI and GPADH moiety to be deeply branched forming monophyletic groups within the TPI and GAPDH clades respectively. It is suggested that there is a mitochondrial (therefore *alpha-proteobacterial*) origin for aerobic respiration in eukaryotes, the genes of which have then been donated to the host nucleus after the endosymbiosis event. Phylogenetic analysis on glycolytic enzymes in the *Stramenopiles* has found them to be more similar to *alpha-proteobacteria* than any other sequenced organism. This goes against the established idea that glycolysis belonged to the host organism, thought now to belong within the Archaea (Der Giezen, 2011; Embley and Martin, 2006; Liaud et al., 2000; Nakayama et al., 2012).

1.4.3 TPI-GAPDH Potential Drug Target?

Whilst the questions about early eukaryotic evolution are interesting and important to our understanding of how life became what it is today, our more pressing need is to discover new drug targets. TPI-GAPDH is a novel enzyme, and has only been found in a handful of *Stramenopiles*, some of which are very important pathogens (*B. hominis, P. infestans*). It makes up two essential enzymes within glycolysis, a cell's primary way of converting glucose into energy, as well as a host of other non-glycolytic functions. This makes TPI-GAPDH a promising target

for chemotherapy. The idea that some *Stramenopiles* have kept TPI-GAPDH as a fusion suggests it could not find a favourable way of splitting them and that the two moieties are potentially dependent on each other either structurally or kinetically for their function. This gives us a potential lead into how it could be utilised as a drug target. Splitting the fusion protein into its enzyme constituents and comparing the function of the split GAPDH with the wild type TPI-GAPDH will show perhaps why the *Stramenopiles* have retained this fusion relic.

1.5 Aims

The discovery of this novel TPI-GAPDH fusion protein in pathogenic protist parasites makes it an exciting area of research. To date, no enzymatic activity characterisation studies have been done on this enzyme. There has been some tentative work into structure but there is little understanding of the functioning of the fusion protein, other than it is assumed to be glycolytically active. There is also no work into looking at TPI-GAPDH as a potential drug target. The aims of this research are:

- To enzymatically characterise the wtTPI-GAPDH enzyme
- To mutate the two moieties of the enzyme to establish the effect they have on each other
- To separate the fusion enzyme into TPI and GAPDH to establish if they structurally depend on each other

1.5.1 Enzymatically characterising wtTPI-GAPDH

This provided the baseline for the rest of the enzymatic assays, as well as demonstrating that TPI-GAPDH is indeed glycolytically active.

1.5.2 Mutating the wtTPI-GAPDH enzyme

This was accomplished by mutating the known active amino acid residues within either the TPI portion of the enzyme or the GAPDH portion by site directed mutagenesis. Enzymatic assays were run over a range of substrate concentrations to establish how the mutations affected the turnover rate and substrate affinity (K_M) of the fusion enzyme. It was expected that mutations in the active site of the TPI portion of the enzyme would cause the apparent rate of GAPDH to increase by increasing the available GAP concentration 22 fold. Allowing for further characterisation of wtTPI-GAPDH. This would also help to establish if the enzymes

kinetically depended on each other or not, as if GAPDH did not function with a TPI mutation then that would potentially suggest a kinetic reason for keeping the enzymes fused.

1.5.3 Separating the enzyme into TPI and GAPDH

Separating the enzymes should establish whether the two moieties of the fusion enzyme depend on each other structurally to function. Previous work has shown that separated GAPDH is not as active as wtTPI-GAPDH, this needed further investigation (data not published). Ideally GAPDH on its own should work faster in assays, as its substrate (GAP) availability would be 22 times higher without the rate-limiting TPI presence. However, the *Stramenopiles* may have kept the fusion enzyme together for a functional reason. If this is the case and the enzymes are less active separately, then there is a reason for further work.

Chapter 2: Materials and Methods

2.1 Production of wild type and mutant proteins

All proteins, except otherwise stated are tagged with a poly-histidine N terminal tag (His-N) for purification. Chemicals are from Fisher unless otherwise stated. Bacterial strains used in this study can be found in Appendix 17.

2.1.1 Competent cells

Preparation of competent cells of BL21 DE3 plysS *Escherichia coli* and XL-1 Blue *E. coli* cell stocks was performed as outlined by (Nishimura et al., 1990).

2.1.2 Transformation of competent cells with wild type plasmid

Wild type TPI-GAPDH PET14b plasmid, containing the gene isolated from *B. hominis* was supplied by Mark Van Der Giezen (University of Exeter). This was transformed into XL-1 Blue cells as outlined by Nishimura *et al*. (1990) and plated onto 0.1mg/ml LB-Ampicillin (Sigma) agar and grown overnight at 37°C.

2.1.3 Plasmid DNA purification

An overnight culture of the XL-Blue stock to be purified was grown in 1ml 0.1mg/ml LB-Amp and centrifuged at 5000g for 1 minute, the supernatant was discarded. The cell pellet was resuspended in 250µl resuspension buffer (30mM

Tris HCl, 10mM EDTA, 100ug/ml Rnase A, pH 8.0), then 250µl of lysis buffer (0.2M NaOH, 1% SDS) and 350µl of neutralisation buffer (4M guanidine hydrochloride, 0.5M potassium acetate, pH 4.2) were gently added to the mix. The mix was centrifuged at 17,900g for 10 minutes and the supernatant was poured into a plasmid miniprep purification spin column (NBSBio). This was centrifuged at 17,900g for 1 minute and the supernatant discarded. 700µl of wash buffer (20mM NaCl, 2mM Tris HCl, 70% ethanol, pH 7.5) was added to the silica column, which was centrifuged for 1 minute, the supernatant discarded then centrifuged for a further minute both at 17,900g. The plasmid was eluted by placing the column in a microcentrifuge tube and adding 50µl of ultra pure water and centrifuging at 11000g for 1 minute. To check presence and concentration of DNA, the DNA was run on a 1% agarose gel containing 1µg/ml ethidium bromide at 70 volts for 30 minutes and visualised under UV light.

2.1.4 Checking for protein expression

Plasmid DNA was transformed as 2.1.2 into a BL21 DE3 plysS stock and two overnight colonies were selected for protein expression. They were grown in 1ml of 0.1mg/ml LB-Amp for 4 hours at 37°C or until A_{600} 0.6 OD is reached, then split and induced one half with 0.1mg/ml IPTG (Melford) to grow at 37°C for a further 4 hours. Protein expression was checked by removing 125µl of cells from each and pelleting cells by centrifugation. Cells were resuspended in 20µl ultra pure water and 20µl of 4x SDS loading buffer (0.5M Tris HCl, 4ml glycerol, 10% SDS, 0.1% bromophenol blue, 1ml mercaptoethanol – for 10ml), incubated for 4 minutes at 95°C and run on a 13% SDS Polyacrylamide gel for one hour at 100 volts. The gel was stained in a 0.25% Coomassie blue G-250 stain (50% methanol, 10% acetic acid) and destained in a 25% methanol, 7% acetic acid destain to visual protein expression. On confirmation of protein expression, a 30% glycerol stock of the protein expressing cells was created and stored in -80°C freezer until needed for large-scale protein purification.

2.1.5 Creation of mutant proteins from wtTPI-GAPDH PET14b by site directed mutagenesis

Three mutants in total were made from wtTPI-GAPDH Pet14b – TPI-GAPDH PET14b C404A (GAPDH mutation), TPI-GAPDH PET14b H96A (TPI mutation) and dTPI-GAPDH PET14b E165A (TPI mutation) by site directed mutagenesis. The location of these mutations can be seen in Appendix 3. This was performed based upon the Quikchange™ protocol (Stratagene).

The reaction mix for mutagenesis contained: 10µl 5x HF buffer (Thermo scientific), 2µl of wtTPI-GAPDH PET14b plasmid, 1.25µl 100ng/µl For primer, 1.25µl 100ng/µl Rev primer (primers for each mutant shown in Table 1), 1µl 10mM dNTP, 1µl Phusion DNA polymerase (Thermo Scientific) and 34.5µl ultra pure water. The mutagenesis mix was amplified using a Biometra personal thermal cycler set to the parameters shown in Table 2.

Table 1: Primers for mutants and amino acid changes

Mutant	Primer
TPI-GAPDH C404A	For 5' GTCGAACGCGTCCGCCACGACGAACTG 3'
	Rev 3' CAGTTCGTCGTGGCGGACGCGTTCGAC 5'
TPI-GAPDH H96A[1]	For 5' GTTGGGCGATCGTGGGCGCCTCCGAGC 3'
	Rev 3' GCTCGGAGGCGCCCACGATCGCCCAAC 5'
TPI-GAPDH E165A	For 5' GTGATCGCCTACGCGCCCATCTGGGCCATTG 3'
	Rev 3' CAATGGCCCAGATGGGCGCGTAGGCGATCAC 5'

[1] Created by K. Favas, MSc student

Table 2: Parameters for site directed mutagenesis with Phusion DNA polymerase

Cycle Step	Temperature (°C)	Time	Number of Cycles
Initial denaturation	98	30 seconds	1
Denaturation Anneal Extension	98 68 72	10 seconds 30 seconds 60 seconds	16-20
Final Extension	72 4	5 minutes hold	1

Once the cycling was complete, 1µl of DpnI (Thermo Scientific) was added and the mix was incubated for 1 hour at 37°C before being transformed into XL-1 Blue *E. coli* competent cells and incubated overnight at 37°C.

Two colonies were selected to purify the newly mutated plasmid as outlined in 2.1.3. The purified mutated plasmids were then sent off for Sanger sequencing to ensuring the mutation had taken.

A further mutant was then created, TPI-GAPDH H96A E165A PET14b, containing both the H96A and E165A mutations using TPI-GAPDH H96A PET14b as a template and the TPI-GAPDH E165A primers.

2.1.6 Analysis of sequence results

The mutated plasmid DNA samples were sent to the DNA Sequencing Facility, Department of Biochemistry, University of Cambridge and Eurofins MWG Operon. The results were returned in a FastA format and analysed in Bioedit (T. Hall, Ibis Biosciences). The mutated sequences were aligned against the known wtTPI-

GAPDH sequence and converted into amino acid sequence to ensure the correct mutation had occurred.

2.1.7 Protein expression in the mutants

The three mutant TPI fusion protein PET14bs were attempted to be transformed into BL21 DE3 plysS but this proved difficult, as they often would not transform, and those that did produced truncated proteins. TPI-GAPDH C404A PET14b transformed into BL21 DE3 plysS with no issue and produced a protein of the correct size, so a glycerol BL21 DE3 plysS stock was created and frozen. A different approach was needed for the mutant TPI fusion proteins.

2.1.8 Ligation of insert into pJC20 vector

The vector was changed to the smaller (2.3 kbp) pJC20 vector (Vector map - Figure 9) to help with transformation efficiency, as well as plasmid purification efficiency compared with the PET14b Vector (4.7 kbp). Two double digest mixes were set up for each mutant TPI fusion protein and wtTPI-GAPDH; one to digest pJC20 and the other to digest the insert out of PET14b. The pJC20 mix contained 16µl of pJC20, 2µl fast digest buffer (Fermentas), 1µl BamHI (Fermentas) and 1µl NcoI (Fermentas). The plasmid insert mix contained 43µl of plasmid DNA, 5µl fast digest buffer, 1µl BamHI and 1µl NcoI. Both mixes were incubated at 37°C for 1 hour then loaded on a 1% agarose gel, ran for 30 minutes at 80 volts and visualised under UV light. Gel bands at 2.3kb for pJC20 and 1.8kb for the insert were cut out and purified by an Eppendorf Perfect Prep Gel Clean Up kit. Purified cut plasmid insert and cut pJC20 vector were then ligated at a 2:1 insert to vector ratio with T4 ligase buffer and T4 ligase (Thermo Scientific) for 30 minutes at 20°C. The newly ligated clones in pJC20s were then transformed into XL-1 Blue, purified by plasmid mini-prep (2.1.3) and transformed into BL21 DE3 plysS for protein expression. On successful transformation and protein expression, glycerol BL21 DE3 plysS stocks were made for all.

2.1.9 Production of poly-histidine C terminus tagged GAPDH mutant

To ensure the position of the His-N tag was not having an effect on the structure of the separated GAPDH (as this is where TPI would join to GAPDH in wtTPI-GAPDH), a version with a poly-histidine C terminal tag (His-C) was made. A 3' primer removing the stop codon from GAPDH and adding a BamHI restriction site and handle (GAPDH His C primer – 5' GCGGGATCCGAGCGATCCACCTTCGCC 3') was calculated using www.yeastgenome.org primer design tool. The primer was designed to have a 55°C annealing temperature. The insert was subcloned from GAPDH PET14b His N by PCR amplification with a reaction mix of 3µl 10pm/µl T7 For primer, 10pm/µl Rev GAPDH His C primer, 2µl 10mM dNTP, 10µl Taq polymerase buffer, 1µl Taq polymerase (New England Biolabs), 1µl GAPDH His N DNA, 80µl ultra pure H_2O. The parameters for PCR amplification are shown in Table 3.

Table 3: The parameters used for PCR amplification of GAPDH His C insert

Cycle Step	Temperature (°C)	Time	Number of Cycles
Initial denaturation	95	2 minutes	1
Denaturation Anneal Extension	95 50 72	30 seconds 30 seconds 90 seconds	30
Final Extension	72 4	10 minutes hold	1

The amplified insert was then run on a 1% agarose gel, visualised under UV light and the band at 1.1kb was cut out and purified by Eppendorf perfect prep gel clean up. A double restriction digest was performed with a reaction mix of 16µl template DNA (either the amplified GAPDH product or pJC20 His C plasmid – vector map in Appendix 11), 2µl BamHI buffer (New England Biolabs), 1µl BamHI (Fermentas), 1µl NdeI (New England Biolabs), incubated at 37°C for 1 hour. The products were

run on a 1% agarose gel, visualised under UV and a band at 2.3kb and 1.1kb for cut pJC20 His C and cut GAPDH insert respectively. Ligation of insert to vector was performed in the same way as outlined in 2.1.8. The GAPDH pJC20 His C was transformed into XL-1 Blue cells, the plasmid DNA was purified by plasmid mini prep outlined in 2.1.3. The purified DNA was then transformed into BL21 DE3 plysS cells as 2.1.2 and protein expression was checked.

2.2 Protein preparation

2.2.1 Overexpression and purification of poly-histidine tagged proteins in BL21 DE3 plysS *E. coli* cells

1L of 0.1mg/ml LB-Amp was inoculated with an overnight culture of desired cell stock and incubated for 3 to 4 hours at 37°C or until A_{600} OD 0.6 is reached. The culture was then induced with 0.1mg/ml IPTG and incubated for a further 3 hours. The culture was centrifuged at 8000rpm for 10 minutes to pellet the cells. Pelleted cells were either stored at -20°C or utilised immediately.

Pelleted cells were thawed on ice and resuspended in 10ml of lysis-equilibration-wash buffer (LEW buffer: 50mM NaH_2PO_4, 300mM NaCl, pH 8.0), the pellet was homogenised and a further 10ml of LEW buffer added. The cells were lysed by sonication on ice for 4 minutes with a minute off between each minute. Lysed cells were centrifuged at 10000g for 15 minutes at 4°C. The supernatant was retained on ice. Protein was purified from the supernatant by gravity flow column chromatography using a Protino Ni-IDA Resin column as per the manufacturers instructions (Macherey-Nagel, 2005). The supernatant was passed through the column 1ml/min and five 2ml fractions collected. The column was then washed with 4ml of LEW buffer, which was also collected in two fractions. Finally to elute, 5ml of elution buffer (LEW buffer, 250mM imidazole, pH 8.0) was added to the column and collected in five 1ml fractions. 0.4g/ml of ammonium sulphate was added to these fractions and they were stored at 4°C.

SDS-PAGE was performed on the fractions to check protein concentration and as a guide to purity.

2.2.2 Protein preparation for spectrophotometer analysis

Ammonium sulphate suspension containing the protein was pelleted by centrifugation at 12000 rpm for 1 minute. The supernatant was removed and the pellet was resuspended in 1ml of 15mM sodium pyrophosphate arsenate buffer (15mM sodium pyrophosphate, 30mM sodium arsenate, pH 8.5). The protein stocks were then centrifuged at 12000rpm for 4 minutes to pellet any insoluble protein, the supernatant kept on ice and the pellet discarded.

2.2.3 Protein concentration after purification

Protein concentration was calculated by measuring the OD of a 1 in 100 dilution of the protein between 250 and 400nm on a spectrophotometer. The final concentration in μM was determined using the extinction coefficient of the protein, the absorbance and the molecular weight. The extinction coefficient and molecular weight of the protein were calculated using the Prot Param tool (http://web.expasy.org/protparam/), the data for each protein can be found in Appendices 12-14.

All of the proteins were adjusted to a working concentration of 15μM for the enzymatic assays.

2.3 Enzymatic Assays

2.3.1 Standard enzymatic assay

Rates of the GAPDH activity of the proteins were established by enzymatic assay. The assay mix contained 520µl 15mM sodium pyrophosphate arsenate buffer, 20µl 7.5mM NAD^+ (Sigma N7004-1G), 20µl 0.1M dithiothreitol (DTT) (Calbiochem) and 20µl 15µM enzyme. A blank rate was established at 340nm on a spectrophotomer and rate assay set to run for 3 minutes collecting data at 0.25-second intervals. Once the spectrophotomer was equilibrated with the assay mix, the assay was started by addition of 20µl 7.5mM DL-GAP (Sigma G5251) or DHAP (Sigma 51269) for (depending on the assay) for a final [S] of 0.25mM after a stable absorbance was established. This was done over a range of concentrations for GAP and DHAP: 0.01mM, 0.03mM, 0.06mM, 0.13mM, 0.25mM and 0.5mM.

Each concentration and enzyme assay was repeated three times.

2.3.2 Pre-equilibration enzymatic assay

Pre-equilibration assays followed the same procedure as 2.3.1, however the assay mix contained 500µl 15mM sodium pyrophosphate buffer, 20µl 7.5mM NAD^+, 20µl 0.1M DTT, 20µl 15mM TPI PET14b enzyme, 20µl 0.25mM GAP. The assays were then started by addition of 20µl 15µM enzyme.

2.3.3 Extra TPI enzymatic assay

Extra TPI enzymatic assays followed the same procedure as 2.3.1, however the enzyme was premixed with 20µl of 15µM TPI PET14b enzyme.

2.4 Size Exclusion Chromatography

0.5mg/ml concentrations of wtTPI-GAPDH, TPI and GAPDH purified protein were prepared in Tris-HCl buffer (20mM Tris-HCl, 100mM NaCl). The samples (25µl) were run on an HPLC at 1ml/min using a Phenomenex Yarra 3u SEC-3000 column. The column was calibrated with bovine thyroglobulin (670kDa), IgA (300kDa), IgG (150 kDa), ovalbumin (44kDa) and myoglobin (17kDa). Molecular mass of the samples was calculated by comparing the elution time against the calibration curve.

Chapter 3: Results

3.1 Production of wild type and mutant proteins

3.1.1 Production of Clones

Four mutants were successfully produced from *B. hominis* wtTPI-GAPDH by site directed mutagenesis, three producing a one amino acid change (TPI-GAPDH C404A, TPI-GAPDH H96A and TPI-GAPDH E165A) and one combining the two TPI amino acid changes to create TPI-GAPDH H96A E165A. The DNA sequence and amino acid changes can be seen in Table 4 and DNA sequences received from the Sanger sequencing results can be found in Appendices 7-10.

All amino acid changes were to an alanine residue. Alanine is the amino acid of choice for mutations as it is the smallest amino acid that retains a side chain - a methyl group (Figure 8). This methyl group replacement eliminates side chain reactivity, whilst not altering the structure or folding of the protein. The mutations chosen in this project were based on well-established active site residues (Appendix 1 and 2). The systematic mutation of residues is essentially performing alanine scanning upon known residues. Alanine scanning substitutes each residue of an enzyme in turn with an alanine to determine the catalytic or functional role of each residue in a protein (Lefevre et al., 1997). Mutating the TPI and GAPDH active residues to alanine should knock out their catalytic functionality in the enzymes, reducing reactivity whilst not affecting the structure.

Table 4: The DNA sequence mutations and amino acid sequence mutations of A) TPI-GAPDH C404A, B) TPI-GAPDH H95A, C) TPI-GAPDH E165A compared to wtTPI-GAPDH

	A) TPI-GAPDH C404A	B) TPI-GAPDH H96A	C) TPI-GAPDH E165A
WT sequence	1246… GCG TCC TGC ACG ACG …1260	322… GTG GGC CAC TCC GAG …336	529… GCC TAC GAG CCC ATC …543
WT amino acid sequence	402 … A S C T T … 406	94… V G H S E …98	163… A Y E P I …167
Mutant sequence	1246… GCG TCC GCC ACG ACG …1260	322… GTG GGC GCC TCC GAG …336	529… GCC TAC GCG CCC ATC …543
Mutant amino acid sequence	402 … A S A T T … 406	94… V G A S E …98	163… A Y A P I …167

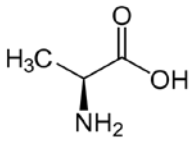


Figure 8: The amino acid alanine

There was difficulty getting the mutated TPI fusion protein plasmids to transform efficiently or express protein in BL21 DE2 plysS. A combination of higher yields of plasmid DNA (the plasmid mini prep described in 2.1.3 provided the highest yields compared to the plasmid mini prep Eppendorf kit, which was used previously) and a change of vector provided the solution. The original PET14b vector was replaced with a pJC20 vector by digesting the insert out of the PET14b vector and ligating it into the pJC20 vector. pJC20 is a smaller vector of 2.3kb with a high copy number, allowing for larger inserts and more efficient transformation (Clos and Brandau, 1994). This led to successful transformation of the mutant TPI fusion clones into BL21 DE3 plysS as well as protein expression.

3.1.2 Protein expression and purification

On successful small-scale expression of all the proteins, large-scale expressions were performed for purification of the protein. Both the PET14b vector and pJC20 vector make use of the T7 promoter system to induce protein production in the cell. The vectors both contain a T7 promoter region and T7 terminator region, within which the gene to be expressed is placed. For example in the pJC20 vector the gene is inserted between the NcoI and BamHI restriction sites (Figure 9). The expression of this gene is dependant on T7 RNA polymerase that is produced on induction of the BL21 DE3 plysS. The addition of IPTG (a non-hydrolysable lactose analogue) initiates the lac operon in the BL21 DE3 plyS to produce T7 RNA polymerase, which transcribe the gene under T7 promoter control in the plasmid vector (Marbach and Bettenbrock, 2012).

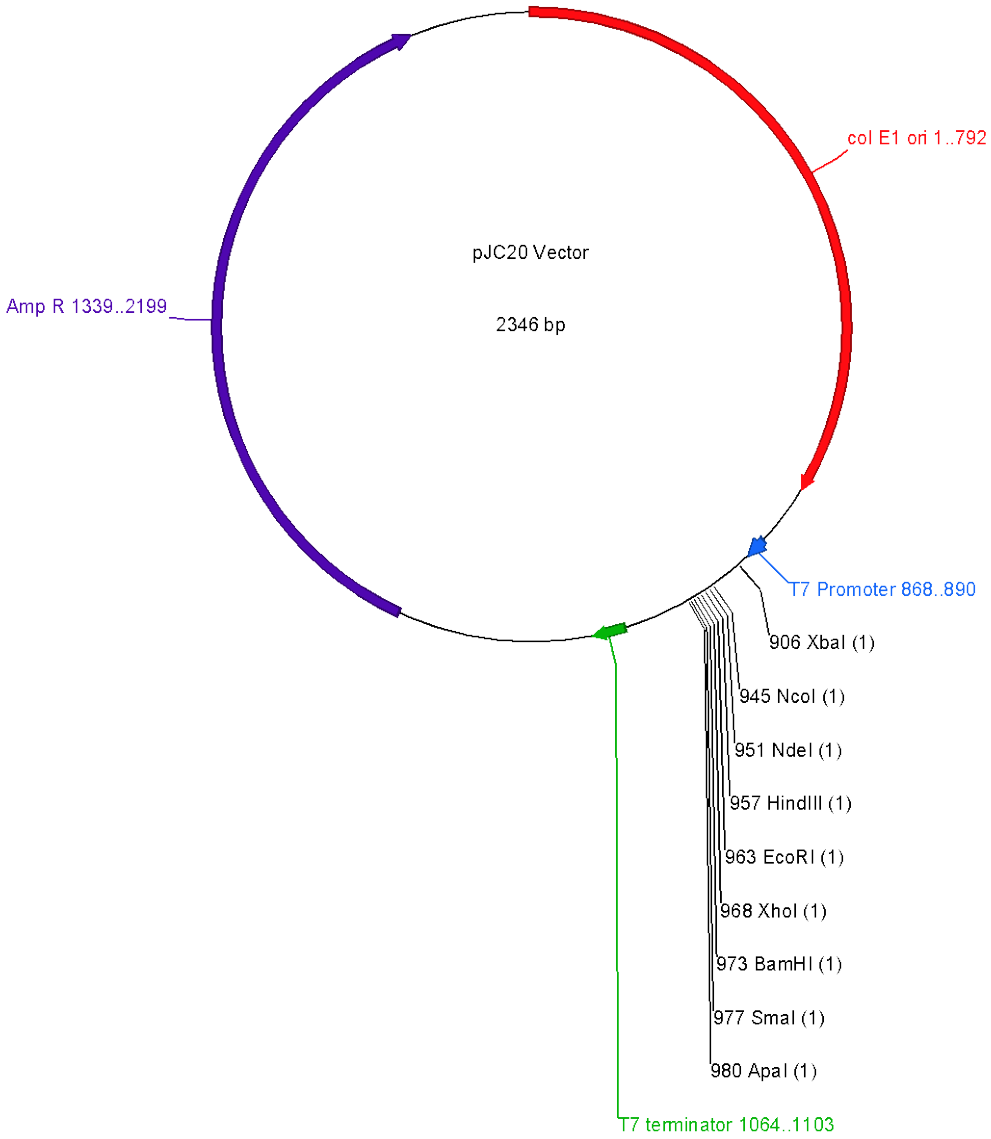


Figure 9: pJC20 vector map showing restriction sites. Gene insert was ligated between NcoI and BamHI.

Protein purification was achieved by gravity flow column chromatography. The proteins were all tagged with a poly-histidine sequence (all on the N terminus except for GAPDH His C, which had a C terminus tag). This allowed them to bind to the Protino Ni-IDA Resin column when the sample was run through. The column contains a dry silica based resin with charged nickel ions (Ni^{2+}), which binds to the polyhistidine tag of the protein and allowing everything else in the sample to flow through into the waste fractions. The column was then washed with LEW buffer to remove anything not bound to it and the proteins were eluted by the addition of an

imidazole elution buffer. Imidazole has a stronger binding affinity to the Ni^{2+} on the columns, replacing the His tagged proteins bound there and allowing them to wash through. The collected fractions were analysed by SDS PAGE, an example of which can be seen in Figure 10, which very clearly shows a large yield of protein in the elution columns in between the 50 kDa marker and the 70 kDa marker, TPI-GAPDH is 65 kDa so the protein was of the correct size to continue on to enzymatic assays.

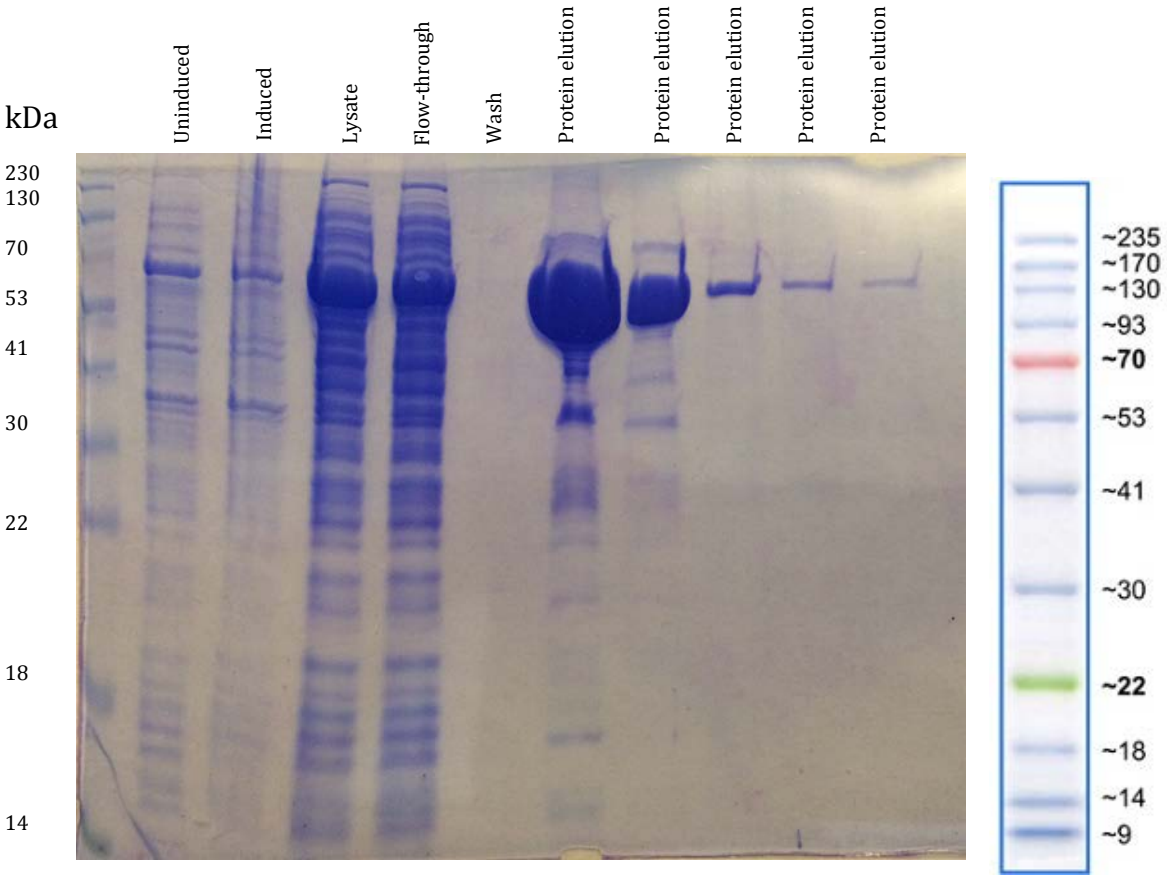


Figure 10: SDS PAGE gel showing protein purification fractions of TPI-GAPDH H96A E165A. Sample = normal cell growth, induced sample = sample induced with IPTG to produce protein, lysate = cell lysate after sonication and centrifugation, flow-through = sample flow through, in this case showing some protein still remaining in the sample, wash = wash with LEW buffer to remove anything unbound, protein elutions = successive protein elutions, showing gradual decrease in eluted protein.

All other proteins were successfully purified using the same protocol. Figure 11 shows a gel of the final purified proteins that were used subsequently. The apparent sizes of the proteins were all of the expected size.

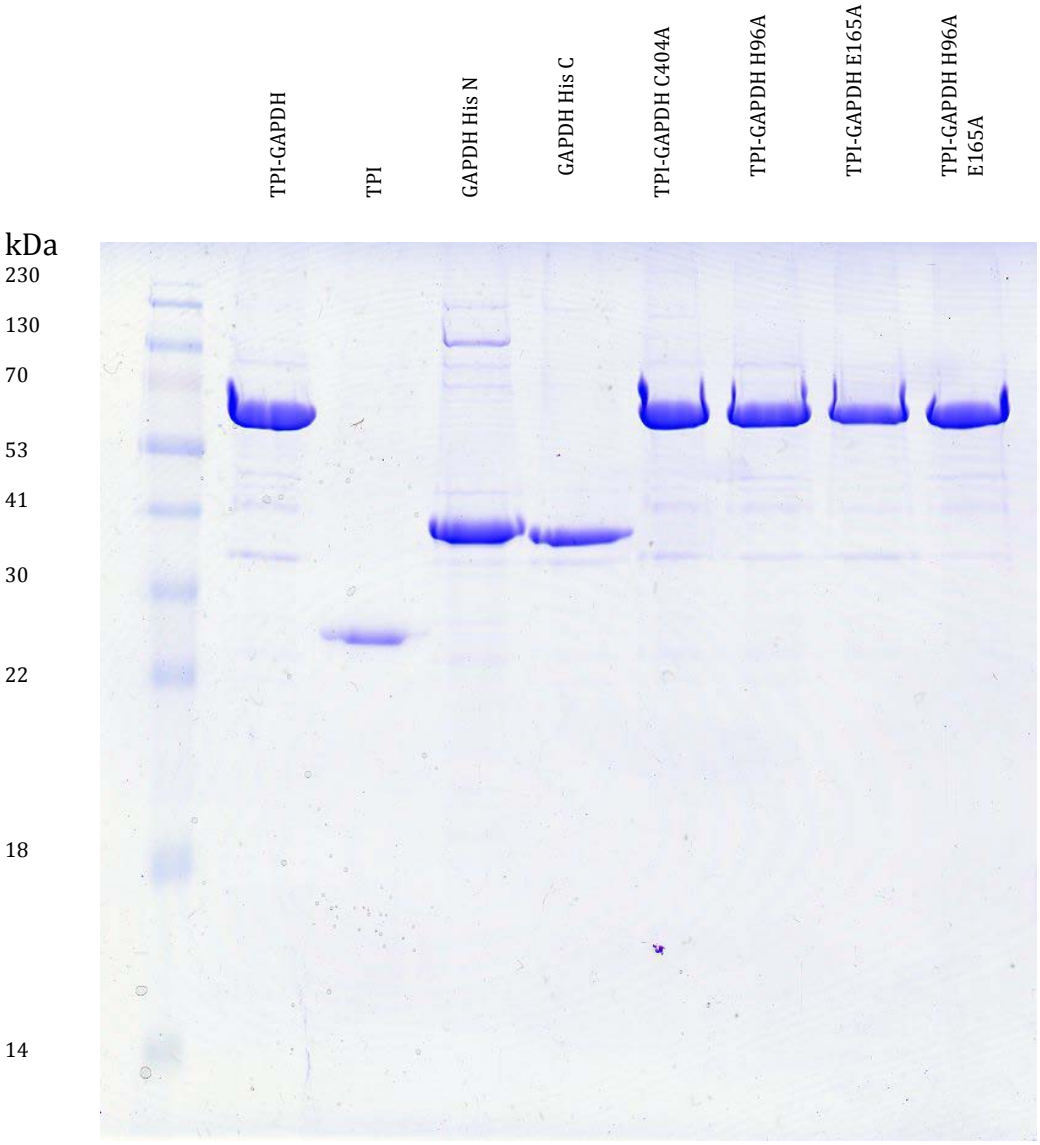


Figure 11: SDS PAGE of re-dissolved ammonium sulphate preparations. Their expected sizes are as follows: TPI-GAPDH = 65kDa, TPI = 28.5kDa, GAPDH His N and His C = 38.7kDa, TPI-GAPDH C404A, TPI-GAPDH H96A, TPI-GAPDH E165A and TPI-GAPDH H96A E165A = 65kDa

3.2 Enzymatic Assays

3.2.1 Concentration of enzymes

The concentration of the enzymes was calculated by diluting the ammonium sulphate precipitated protein stocks in 15mM pyrophosphate arsenate buffer used in the enzymatic assays. The absorbance of a 1 in 100 dilution of the protein was measured between 250 to 400 nm (Figure 12).

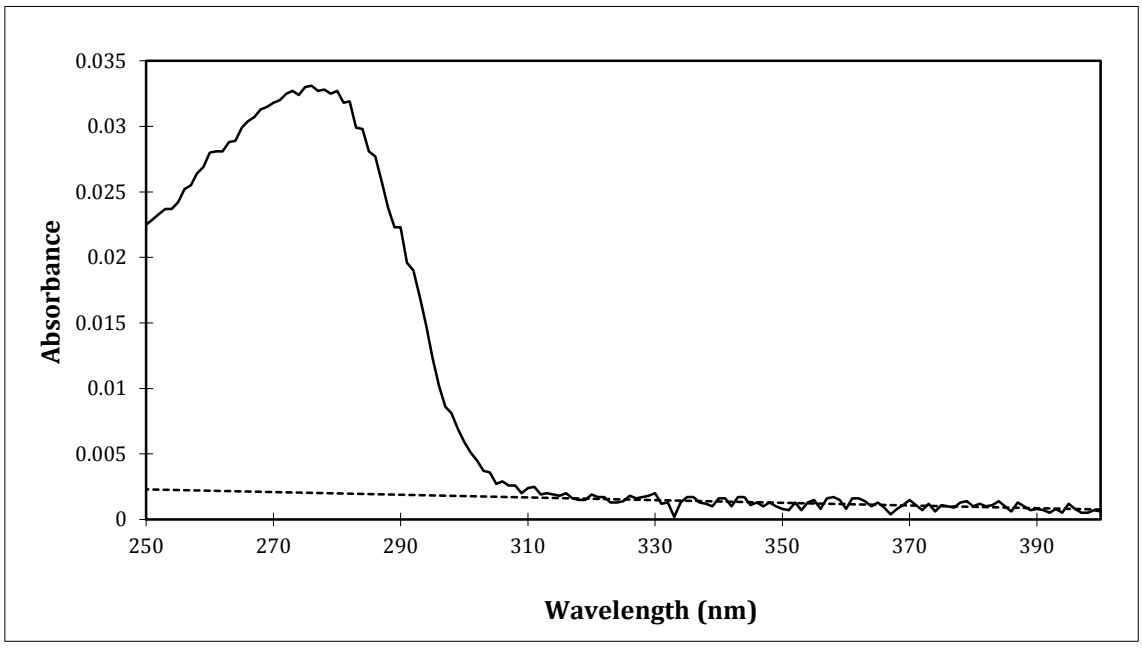


Figure 12: Absorbance spectrum of TPI-GAPDH H96A E165A between 250nm and 400 nm, showing a peak at 280nm. Amino acids with aromatic rings in their side chain absorb light at 280nm, the level of absorbance at 280nm is therefore specific to the number of aromatic residues a protein has.

The concentration of the proteins was calculated by the following equation:

$$C = \frac{A \times dil \times \varepsilon}{mwt}$$

Where C= concentration in mM, A = absorbance at 280nm, dil = dilution of 1 in 100, ε = extinction coefficient and mwt = molecular weight (as predicted by Prot Param)

All protein concentrations were adjusted to 15µM for use in the assays. However, there is a potential that NAD^+ can come bound to GAPDH as it is a co-enzyme and this needs to be taken into account in regards to the concentration calculations. From an absorbance trace of NAD calculated based upon an extinction coefficient of 16900 $M^{-1}\,cm^{-1}$ at 260, the extinction coefficient at 280 is 5000 $M^{-1}\,cm^{-1}$. The extinction coefficient of TPI-GAPDH is 81820 $M^{-1}\,cm^{-1}$, therefore if a maximum of one NAD^+ molecule is bound to one GADPDH is assumed the maximum error in the protein concentrations would be (5000/81820) 6%.

3.2.2 Enzyme Kinetics – Establishing enzyme rates

The rate of a reaction depends on a number of factors: concentration of enzyme [E], concentration of substrate [S], the maximum rate of the enzyme V_{max}, and the Michaelis Constant K_M (the [S] required for the enzyme to run at half V_{max}). The assays used were designed to generate pseudo first order conditions to allow Michaelis-Menten (M-M) plots to be generated. To make the assays first order in respect to the assayed substrate, two conditions have to be satisfied.

Firstly the [E] used was the same in all assays. This means it can effectively be ignored and the reaction considered pseudo-first order with respect to [S]. It should be noted that any variations in [E] will effect the measured rate as the reaction is really higher order.

Secondly the concentration of other substrates, in this case NAD^+ is in excess so it effectively stays constant in the initial part of the reaction from which the rate is determined.

Under these conditions, the variation in measured rate (V_0) as [S] changes allows us to calculate the V_{max} and K_M of the enzymes. V_0 is dependant on [S], as described by the first order M-M equation where V_0 = initial rate:

$$V_0 = \frac{V_{max}[S]}{K_m + [S]}$$

Therefore varying [S] will allow us to establish a set of observed rates that can be used to determine the specific V_{max} and K_M of an enzyme. Figure 13 shows a typical M-M plot of rate against [S] demonstrating V_{max} and K_M, as well as the linear period (when rate is effectively directly proportional to [S]) (Lesk, 2010). When in the linear region of the M-M plot increasing the [S] will increase the rate proportionately. Using Figure 13 as an example, it can be seen that increasing the [S] from 0 to 1, will increase the rate from 0 to 1. However, the closer the [S] gets to K_M and V_{max} the lower the difference in rate will be. A [S] increase between 4 and 8 equates to a rate difference of about 0.5 rather than 4.

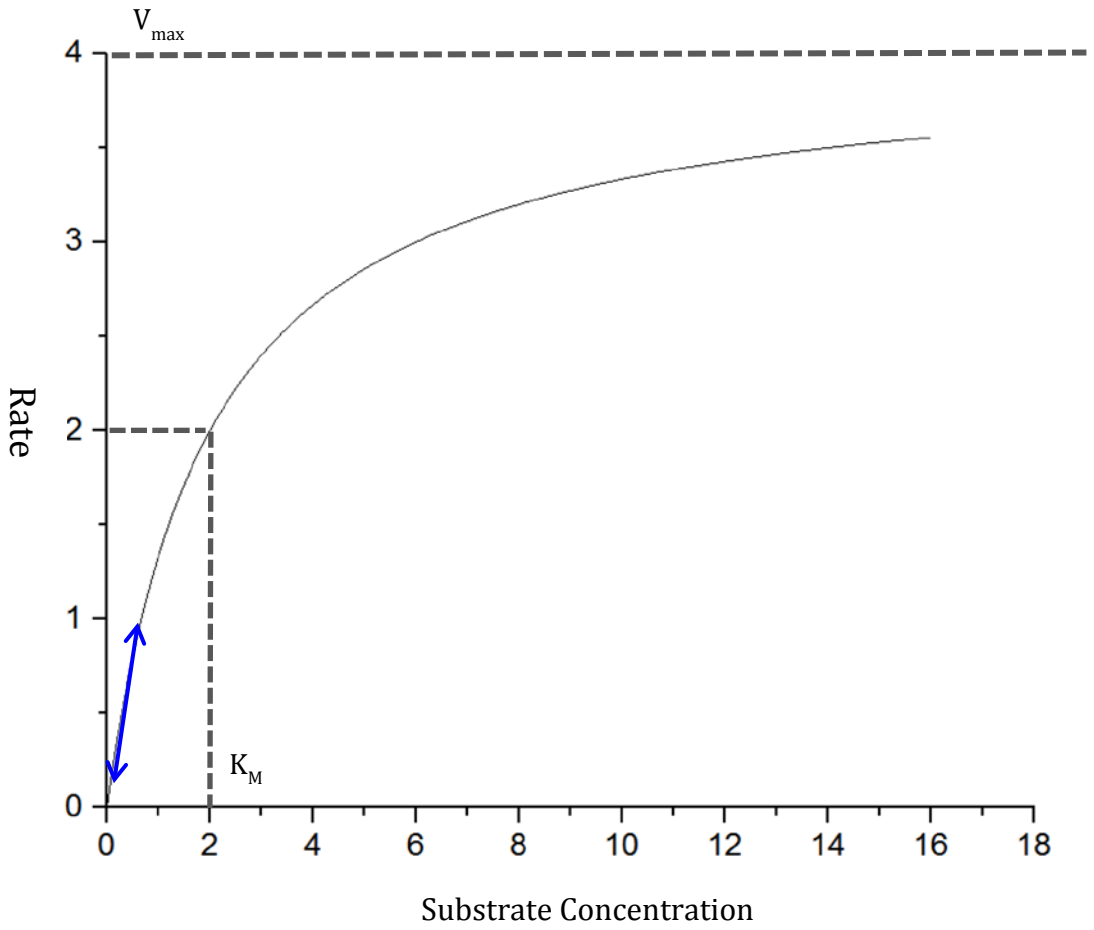


Figure 13: An example Michaelis-Menten curve showing Rate against [S], with a V_{max} of 4 and a K_M of 2 ([S] to reach half Vmax). The linear region of the enzyme is also illustrated (blue arrows), where rate is approximately proportional to [S].

The enzymatic assays gave a reading of absorbance at 340nm over time; the absorbance at 340 nm is due to the production of NADH (which absorbs light at 340nm) as a product of the dehydrogenase activity of GAPDH, making it a convenient measure of the activity of the enzyme. The absorbance against time curve was converted into a rate by measuring the slope of the curve during the initial part of the reaction when it is still in the linear region of the curve (Figure 14). All initial rate data can be found in Appendix 15.

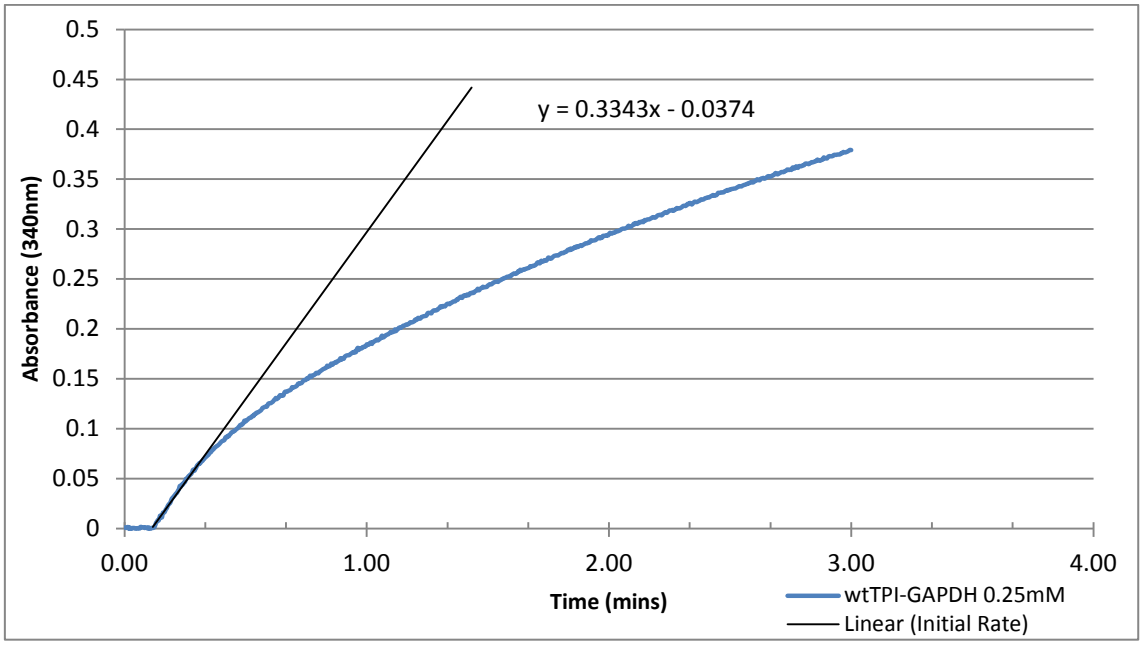


Figure 14: Absorbance of NADH produced over time from wtTPI-GAPDH with 0.25mM of GAP as substrate. The rate was established by calculating the initial rate of reaction slope as shown. VisionPro spectrophotometer software was used to calculate all initial rates.

Before rates were plotted against [S], it could be easily visualised that increasing the [S] increased the rate of reaction, as in Figure 15. The progressive increase of the absorbance at the end of the 3 minutes shows that the [S] used were still in the lower part of the M-M curve, where the rates and amounts of product produced are still increasing as [S] increases (see Figure 13).

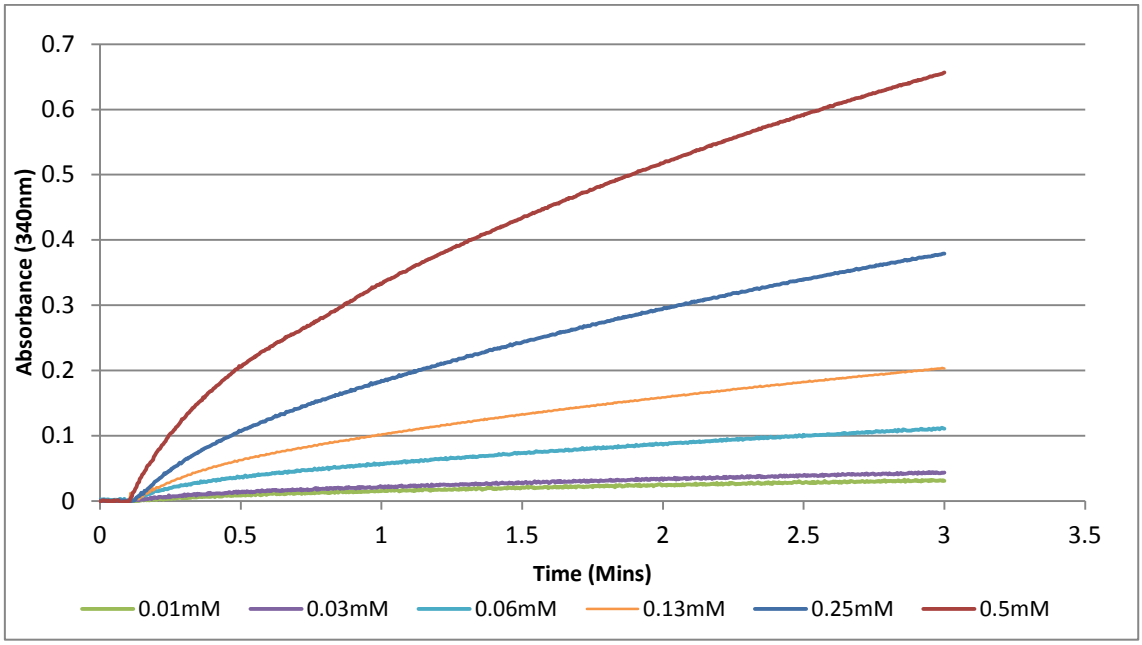


Figure 15: Absorbance of NADH at 340nm of over time from wtTPI-GAPDH over a range of concentrations of GAP as substrate

Differences between enzymes could also be preliminarily identified by the raw results gained from the enzymatic assays as demonstrated in Figure 16. This shows the rates of different enzymes at the same [S]. It can be seen that ddTPI-GAPDH was producing NADH at a much faster rate than wtTPI-GAPDH and faster still than GAPDH His C. Comparatively GAPDH His N was producing very little NADH.

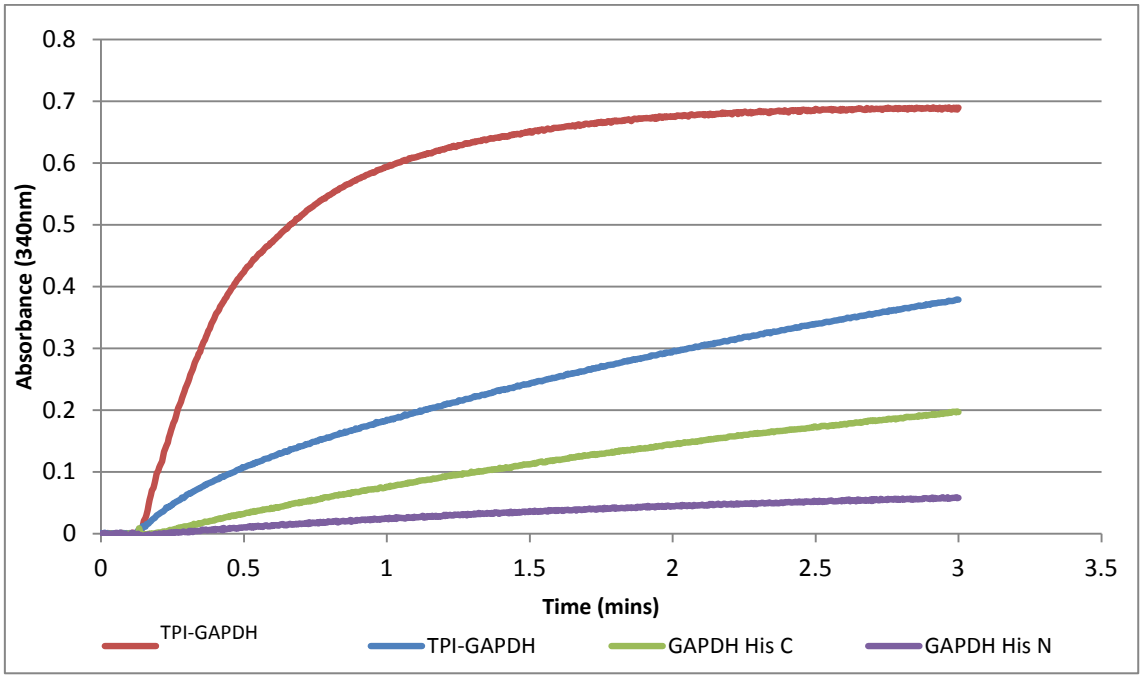


Figure 16: Absorbance of NADH over time from TPI-GAPDH H96A E165A, wtTPI-GAPDH, GAPDH His C and GAPDH His N with 0.25mM GAP as substrate. wtTPI-GAPDH has a 2 fold increase in NADH production over GAPDH His C and TPI-GAPDH H96A E165A has a further 1.5 fold increase over wtTPI-GAPDH

3.2.3 Calculating V_{max} and K_M

To calculate the V_{max} and K_M of each enzyme, the average rates were plotted Origin 9.1 (OriginLab) to curve fit the data to give V_{max} and K_M based on the M-M equation. This also allowed for a fitted curve that extends past the range of the data to be produced as seen in Figure 17. The black squares within the pink arrows show the actual data collected, the rest of the line represents the fitted curve based on the V_{max} (P1) and K_M (P2) values calculated from simulating the raw data.

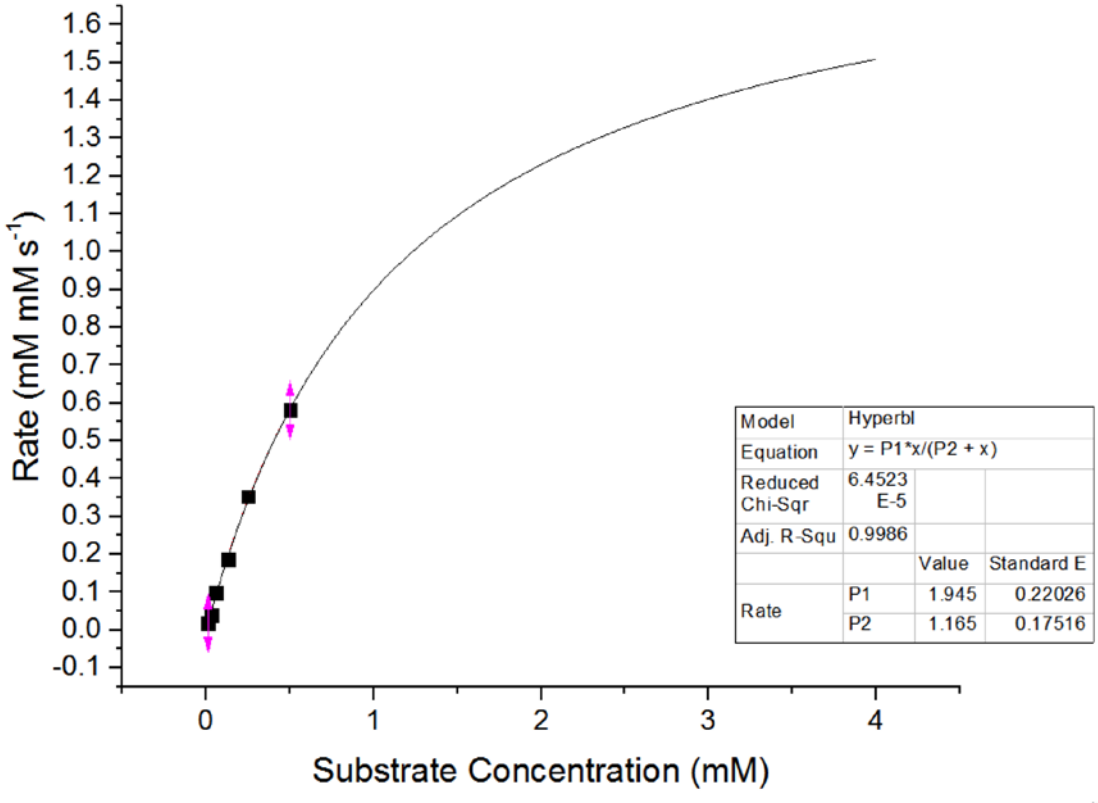


Figure 17: Fitted curve of wtTPI-GAPDH with GAP as the substrate. The black squares within the pink arrows represent the average rates of wtTPI-GAPDH over a range of substrate concentrations. The V_{max} (P1) and K_M (P2) were calculated by simulating the rates. The curve was fitted using the V_{max} and K_M values to extend past the inputted data.

As can be seen in Table 5, the V_{max} varies, when ideally the V_{max} for the enzymes should be the same as the assays were all measuring an unchanged GAPDH activity. This variation makes the true K_M of the GAPDH portion of the enzyme difficult to determine, as there is some interdependence between variation in V_{max} and K_M. However, there is a general trend that the K_M for the mutated TPI fusion enzymes is lower than the others. A lower K_M tells us less substrate is required to reach half V_{max}. Therefore the lower the K_M, the higher the rate of the enzyme at the same [S], until close to V_{max}. GAPDH does not function as well when separated from TPI, as seen by the higher K_M and lower V_{max}.

Table 5: V_{max} and K_M values for all enzymes with GAP as the substrate.

Enzyme	$V_{max\ (mM\ mM^{-1}\ s^{-1})}$	K_M (mM)
wtTPI-GAPDH	1.95	1.17
GAPDH His N	0.07	0.2
GAPDH His C	0.92	2.09
TPI-GAPDH C404A	0.02	0.13
TPI-GAPDH H96A	3.28	0.47
TPI-GAPDH E165A	5.42	1.11
TPI-GAPDH H96A E165A	4.53	0.6

Differences in V_{max} make the results difficult to determine, however mutations in the TPI portion of the enzyme generally have a lower apparent K_M. GAPDH His N and TPI-GAPDH C404A both had very low rates even at higher [S], making the calculated V_{max} and K_M results unreliable. GAPDH His C has a 1.8 times higher K_M than wtTPI-GAPDH.

The V_{max} should have been the same for the wt and mutant TPI fusion enzymes as GAPDH activity is being measured and only the TPI portion has changed, so any effect of that should be on the K_M. Therefore the curve fitting was repeated, this time however the V_{max} was constrained so that the enzymes analysed had the same V_{max}. This showed a much clearer pattern in K_M between the TPI mutations (Table 6). However, this only gives an indication of what the K_Ms of GAPDH would be under the assumption that there are no other factors influencing the enzymes, potentially the TPI mutations could be having an effect on the structure and folding of the enzyme that could lead to a change in the V_{max}. Therefore these K_M values with constrained V_{max}s are not absolute.

Table 6: K_Ms of wt and mutant TPI fusion enzymes with GAP as substrate when V_{max} is constrained.

Enzyme	V_{max} (mM mM^{-1} s^{-1})	K_m (mM)
wtTPI-GAPDH	4.22	3
TPI-GAPDH H96A	4.22	0.71
TPI-GAPDH E165A	4.22	0.79
TPI-GAPDH H96A E165A	4.22	0.53

TPI-GAPDH H96A E165A has the lowest K_M, a 5.6 fold difference to wtTPI-GAPDH when the V_{max} is constrained, showing that TPI-GAPDH H96A E165A is most effective at reducing TPI activity as apparent GAPDH K_M is decreased the most.

All of the mutant TPI fusion proteins had a significantly lower apparent K_M than the wtTPI-GAPDH, with TPI-GAPDH H96A E165A having a 5.6 times lower K_M than wtTPI-GAPDH. Showing that the apparent rate of GAPDH is higher when TPI is mutated. The rate is proportionately higher to the reduction in apparent K_M in the mutant TPI fusion proteins compared to wtTPI-GAPDH, as seen in the fitted curves in Figure 18. Extending the fitted curves to higher [S] puts the V_{max} and K_M into context when comparing between enzymes. It also means that the shape of the M-M plot for a particular enzyme can be visualised without a full set of data.

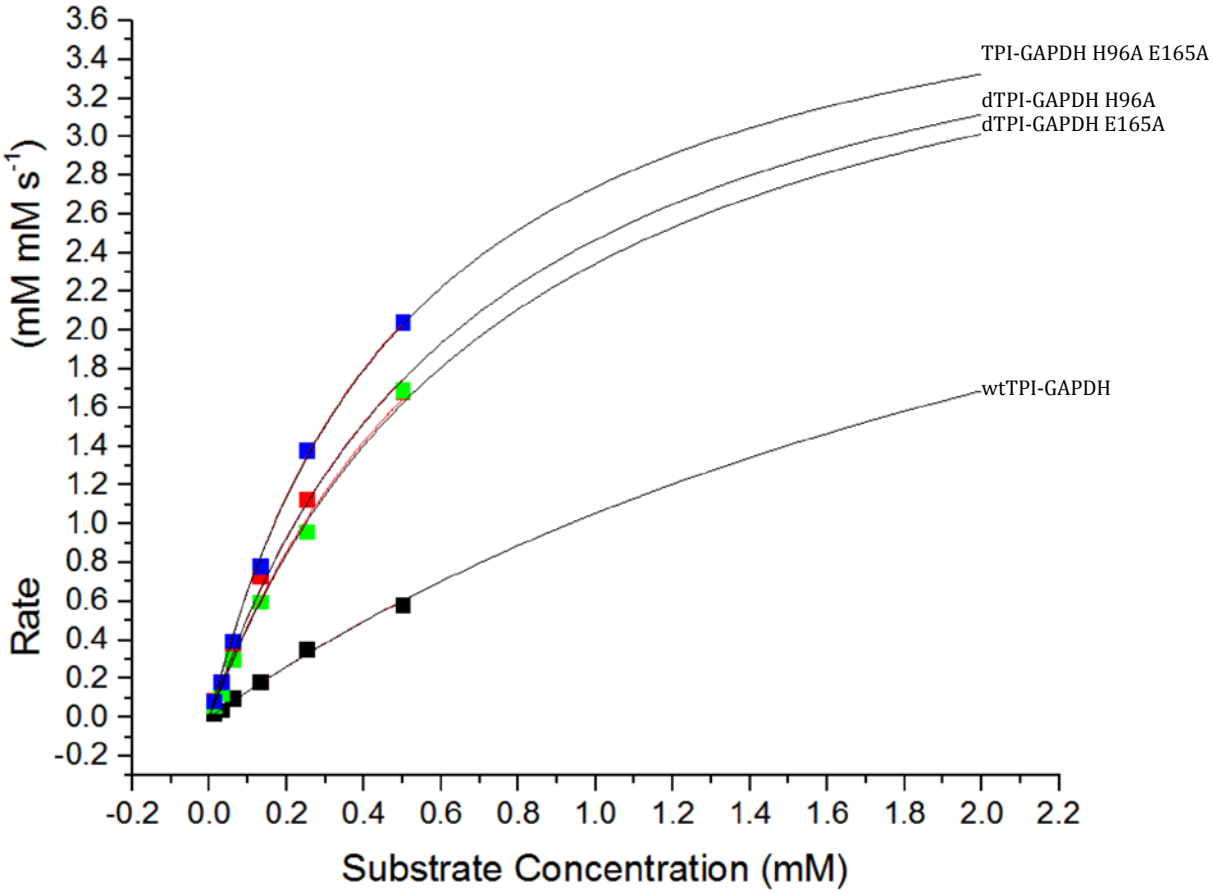


Figure 18: Simulated curve fit of wtTPI-GAPDH and the mutant TPI fusion proteins with constrained V_{max}. TPI-GAPDH H96A E165A (blue), TPI-GAPDH H96A (red) and TPI-GAPDH E165A (green) have significantly faster rates than wtTPI-GAPDH (black), the difference proportional to a lowering in their K_Ms.

The TPI-GAPDH C404A, GAPDH His C and GAPDH His N cannot be compared in the same way the mutant TPI fusion proteins can be. Separating GAPDH from TPI removes the competitive effect of TPI from the enzyme, so a change in both apparent V_{max} and K_M is expected. Table 5 shows GAPDH His C to have a 1.8 times higher K_M than wtTPI-GAPDH. This was unexpected as when separated from TPI, GAPDH should be able to perform a maximum of 22 times more efficiently than the wt enzyme based on where the [S] is on the M-M curve (Figure 13). GAPDH His N and TPI-GAPDH C404A both had low rates even at higher concentrations (below 0.05 mM mM^{-1} s^{-1} at 0.5 mM GAP compared to 0.5 mM mM^{-1} s^{-1} wtTPI-GAPDH), giving an artificially low V_{max} and an unrealistic K_M.

3.2.4 Pre-equilibrating the enzymatic assays with TPI shows TPI competes with GAPDH for GAP

To model a pre-equilibrated reaction, as is generally assumed to be true in a cellular environment, TPI was added to the assay mix with GAP before the assay began and the enzyme being assayed was added. This meant the effect of TPI was separated from the catalysis by GAPDH as it allowed for the equilibrated 22:1 DHAP to GAP ratio to form before the reaction was started.

The results from these assays highlighted some interesting characteristics (Table 7; Figure 19). There was variation in the rates between the wtTPI-GAPDH and mutant TPI fusion enzymes in the pre-incubated assays (red), although these differences were comparatively small (less than a factor of 2).

Table 7: A comparison between the standard enzymatic assay, pre-equilibrating the assay mix with TPI and adding extra TPI to the enzyme being assayed.

Enzyme	Control Assay Rate ($mMmM^{-1}s^{-1}$)	Pre-equilibrated with TPI ($mMmM^{-1}s^{-1}$)	Extra TPI ($mMmM^{-1}s^{-1}$)	Pre-equilibrated with TPI : Normal rate
wtTPI-GAPDH	0.35	0.16	0.19	1 : 2.19
TPI-GAPDH E165A	0.99	0.07	0.12	1 : 13.6
TPI-GAPDH H96A	1.13	0.13	0.55	1 : 8.64
TPI-GAPDH H96A E165A	1.35	0.12	0.14	1 : 11.7
GAPDH His C	0.09	0.01	0.01	1 : 18.6
GAPDH His N	0.04	0.003		1 : 11

Pre-equilibrating the assay mix with TPI reduced the rate for all enzymes. Only GAPDH His C has a rate reduction close to the maximum possible 22 fold reduction. The difference between the wtTPI-GAPDH normal rate and pre-incubated with TPI rate suggest the TPI portion of the enzyme is competing against GAPDH for GAP.

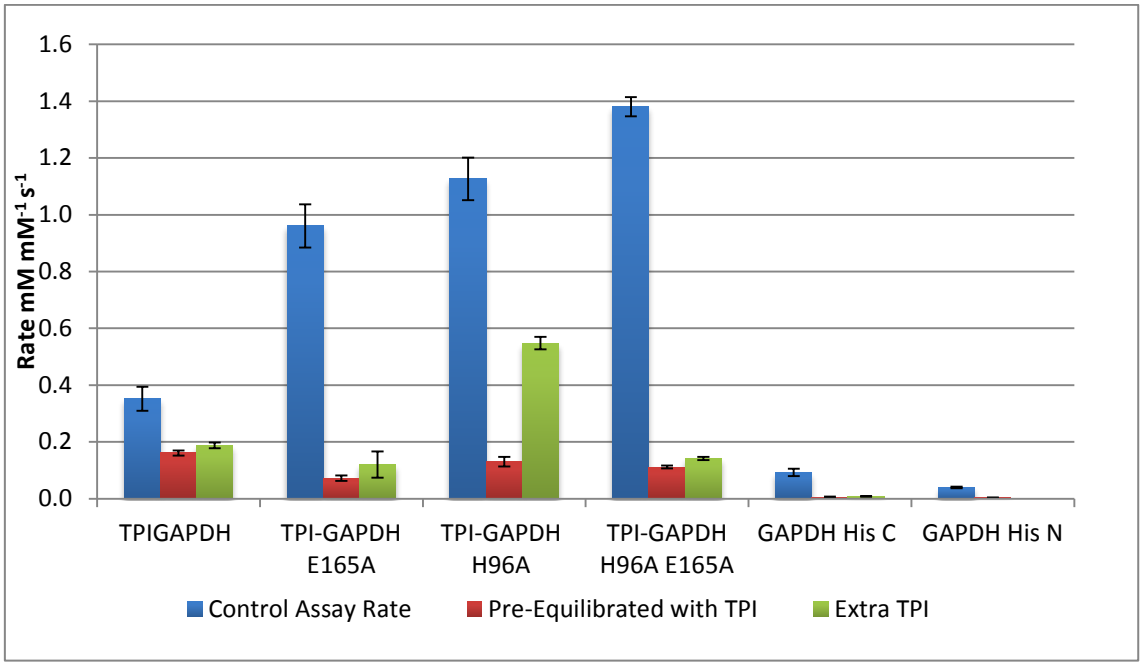


Figure 19: Rates of enzymatic assays with 0.25mM GAP as substrate. The mutant TPI fusion proteins provide a much faster apparent rate in GAPDH as it has removed some of TPIs ability to compete against GAPDH for GAP. wtTPI-GAPDH rate is halved when the assay mix is pre-equilibrated, rather than being the same demonstrating that TPI is competing with GAPDH for GAP in TPI-GAPDH.

Pre-equilibrating the wtTPI-GAPDH assay with TPI decreased the rate by 2 fold compared to the control assay (blue). Pre-mixing wtTPI-GAPDH with extra TPI (green) also decreases the apparent rate by slightly less than two fold (Figure 19), showing that adding more TPI decreases the rate of TPI-GAPDH but not as much as pre-equilibrating with TPI.

The mutant TPI fusion proteins showed a significant decrease in rate when pre-equilibrated with TPI, with TPI-GAPDH E165A showing 13 fold decrease in rate compared with the control assay rather than TPI-GAPDH H96A E165A (which has the lowest K_M, so should have been the most effective mutation, however TPI-GAPDH E165A had a lower average rate with the pre-equilibrated TPI assay, this has been reflected in the normalised data).

Pre-incubating GAPDH His C with TPI showed an 18-fold difference compared to the control assay enzyme, close to the maximum 22-fold effect on available substrate at equilibrium. GAPDH His C has a 32 times slower apparent rate than wtTPI-GAPDH during the pre-equilibrated assays, compared to being 3.2 times slower in the control assays. This shows that the separated *B. hominis* GAPDH His C is a less active enzyme when not attached to TPI, as this comparison to wtTPI-GAPDH in the pre-equilibrated assay shows. This is an effect both on K_M and V_{max}.

Table 8: K_M of wtTPI-GAPDH and TPI mutated enzymes with 0.25mM DHAP as substrate when V_{max} is constrained.

Enzyme	V_{max}	K_M
wtTPI-GAPDH	1.05	0.52
TPI-GAPDH H96A	1.05	25.59
TPI-GAPDH E165A	1.05	34.80
TPI-GAPDH H96A E165A	1.05	40.00

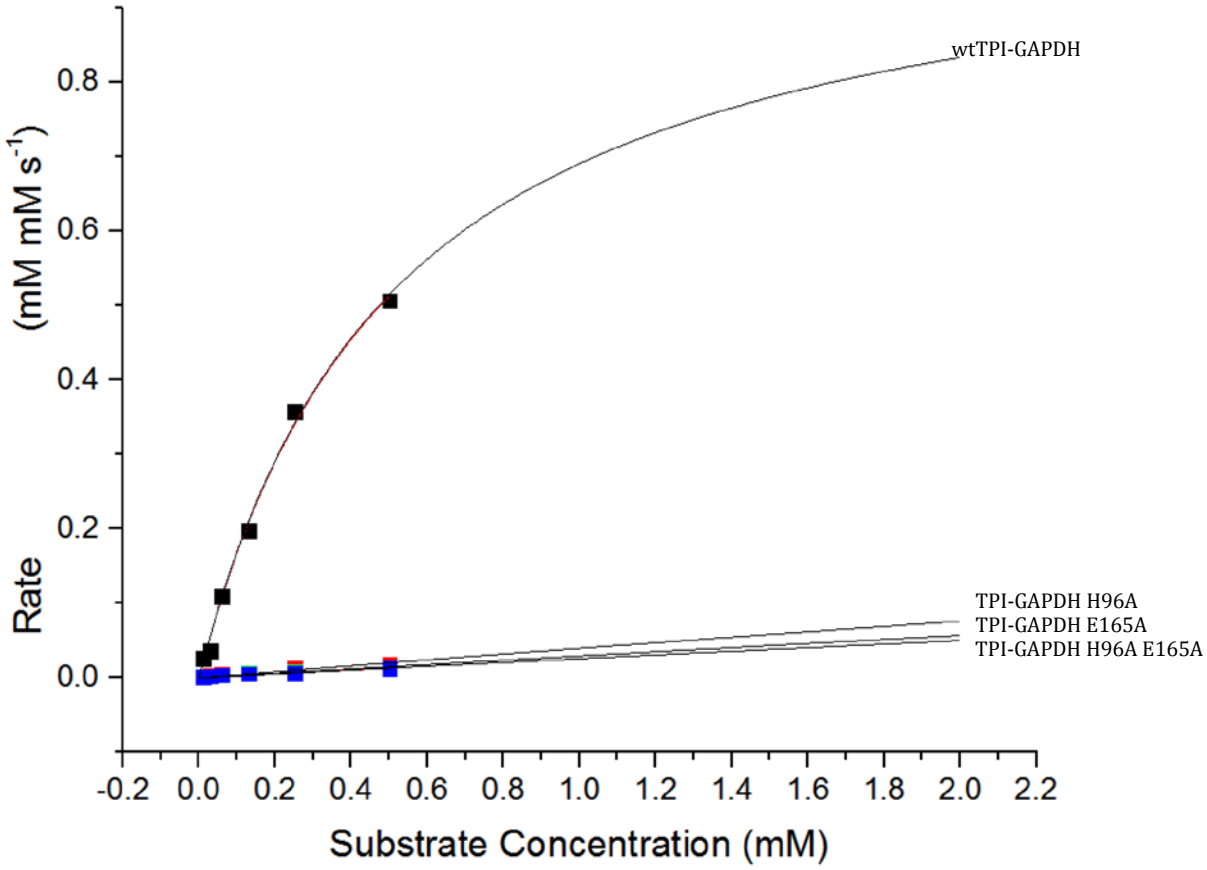


Figure 20: Fitted curve of wtTPI-GAPDH and the mutant TPI fusion proteins to simulated K_M with constrained V_{max} and DHAP as substrate. Compared to wtTPI-GAPDH the mutated TPI enzymes have a much slower rate than wtTPI-GAPDH with DHAP but still with some activity, even TPI-GAPDH h96a e165a is not a complete knockout of the enzyme.

3.3 Size Exclusion Chromatography

Size exclusion chromatography gave preliminary data to suggest that TPI-GAPDH is a dimer of TPI and GAPDH dimers. The size exclusion column was calibrated with a selection of proteins with a known size and their elution times were used to produce a calibration curve of log molecular weight against elution time. When the samples were chromatographed the elution times could then be used with the calibration curve to calculate the molecular weight of the sample, as larger molecules will elute first (Figure 21).

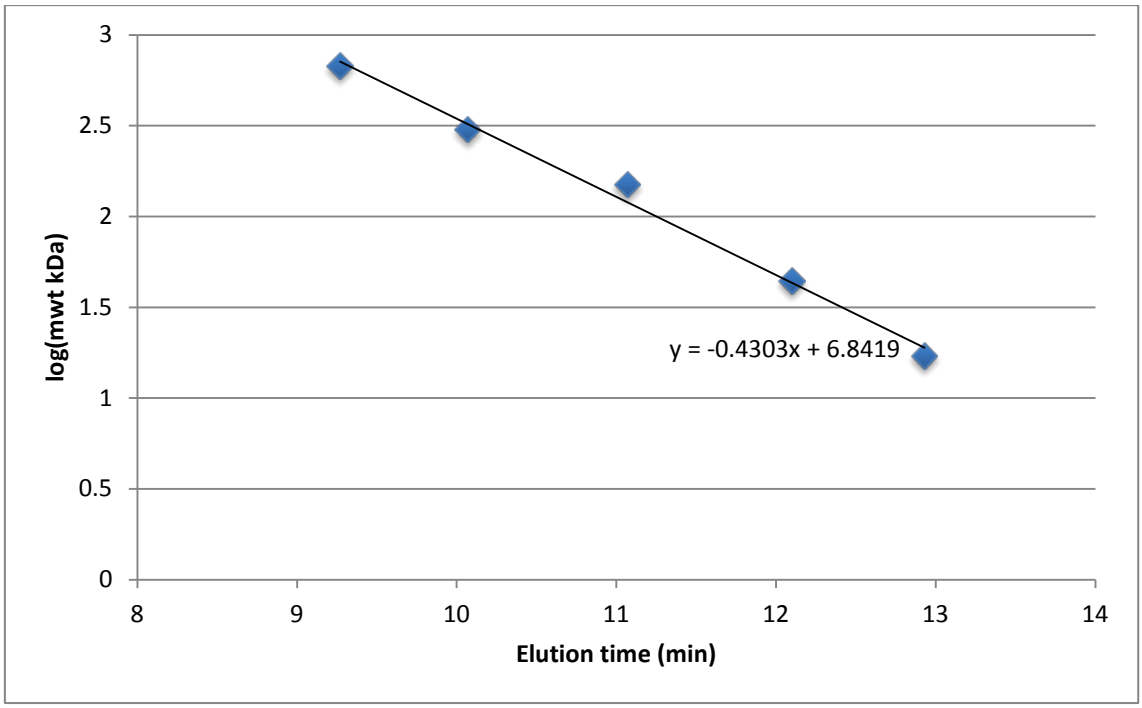


Figure 21: Gel filtration calibration curve of log(mwt kDa) against elution time from the size exclusion column. The data points represent the proteins used for calibration as outlined in 2.4.

Figure 22 shows the elution peaks over time. The calculated molecular weight for a TPI subunit is 28kDa and monomeric GAPDH is 37kDa (Appendix 13, 14). The peaks indicated by the arrow at around 11-12 minutes therefore appear to be monomeric wtTPI-GAPDH (blue) with an estimated mwt of 64kDa, dimeric TPI (green) 58kDa and dimeric GAPDH (red) 64kDa. The dimeric GAPDH mwt is a little less than expected (74kDa). The first wtTPI-GAPDH peak to the left of the arrow is indicative of a "dimer of dimers" TPI-GAPDH with a mwt of 141kDa. However, the smaller peak to the left of that one would suggest a tetrameric structure for GAPDH with two TPI dimers attached with a molecular weight of 232kDa. Elution times and calculated molecular weights can be seen in Table 9.

There does appear to be a small tetrameric TPI peak to the left of the arrow at around 11 minutes, suggesting the TPI has formed aggregates, this could potentially be true of the tetrameric TPI-GAPDH peak if the proteins samples are

not in suitable conditions. The large monomeric TPI-GAPDH peak could also be explained by unfavourable conditions causing the enzyme has dissociated into its monomers.

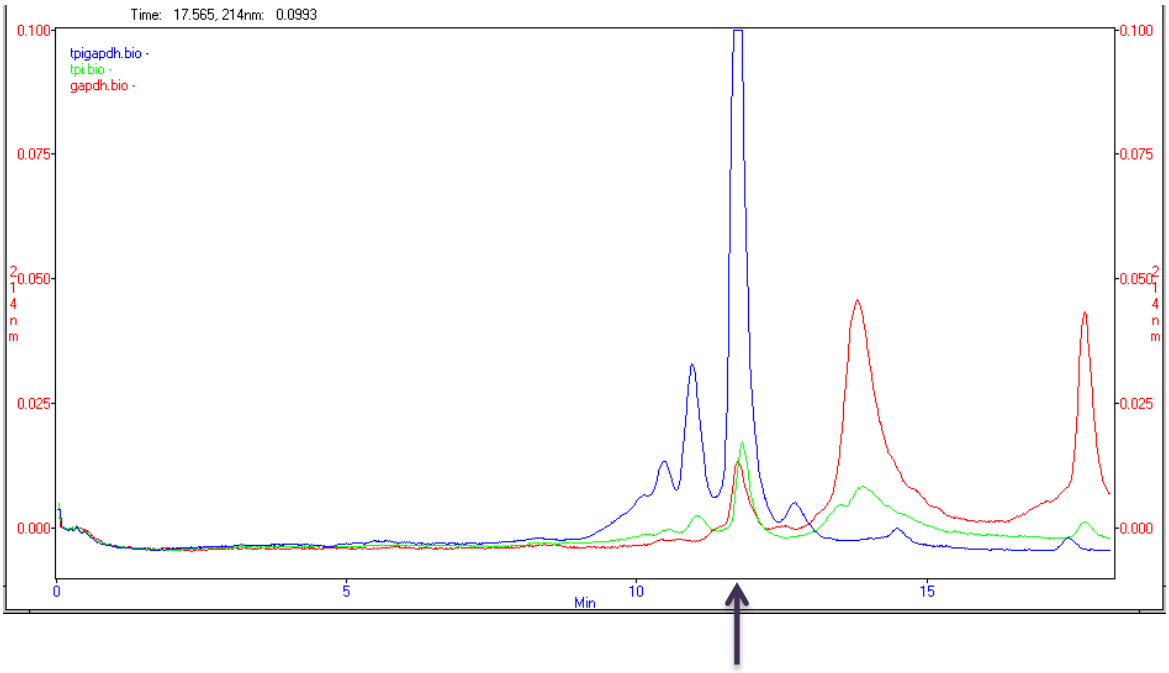


Figure 22: Chromatograph of size exclusion showing elution peaks over time of wtTPI-GAPDH (blue), TPI (green) and GAPDH (red). Arrow highlights dimeric TPI and GAPDH peaks.

Table 9: Molecular weights of preliminarily identified samples based on the elution times of the samples against using the calibration curve in Figure 21.

Preliminary Identification	Elution Time of Peak (mins)	Molecular Weight (kDa)
Dimeric TPI	11.8	57.9
Dimeric GAPDH	11.7	63.9
Monomeric TPI-GAPDH	11.7	63.9
Dimeric TPI-GAPDH	10.9	141
Tetrameric TPI-GAPDH	10.4	232

Chapter 4: Discussion

4.1 The TPI-GAPDH fusion enzyme performs the same glycolytic function as in other organisms.

As expected based on the literature (Liaud et al., 2000), the wtTPI-GAPDH fusion enzyme from *B. hominis* behaves in the same way expected for glycolytic function as the separate enzymes found in other eukaryotes. During glycolysis, TPI interconverts DHAP and GAP, maintaining the thermodynamic 22:1 ratio within the cell. This function was present in the wtTPI-GAPDH fusion enzyme as demonstrated by the fusion enzyme's ability to function when DHAP was used as a substrate (Figure 20, Table 8). The mutated TPI fusion enzymes also showed that TPI's normal glycolytic function is decreased by the mutations, as their activity with DHAP is much lower and there is an apparent increase in GAPDH rate measured using GAP as substrate when mutations are present. This is due to a decrease in the rate at which TPI converts the added GAP to DHAP, increasing the available substrate concentration for GAPDH. Similarly the GAPDH portion of the fusion enzyme performs the same glycolytic function as it would in cells, however with a slightly higher K_M than other organisms. For example the *B. hominis* wtTPI-GAPDH had a K_M of 1.17mM, whilst *T. cruzi* GAPDH has been found to have a K_M of 0.5mM and *Staphylococcus aureus* has a slightly lower K_M of 0.31mM (Cardoso et al., 2008; Mukherjee et al., 2010). There is the potential that the addition of the TPI moiety attached to the N-terminus of the fusion enzyme causes some structural changes that reduces the function of GAPDH.

4.2 Knocking out the TPI moiety in the fusion enzyme causes the apparent K_M of GAPDH to decrease but not by as much as predicted.

The mutations inserted into the TPI moiety of TPI-GAPDH were anticipated to cause a maximum of a 22-fold increase in GAPDH rate and 22-fold reduction in apparent K_M depending on the [S]. If [S] is significantly below K_M in the linear region of the M-M plot (Figure 13), then around a 22 fold difference in rate was expected, though once the [S] rises towards K_M and V_{max} the difference in rate between wtTPI-GAPDH and the mutant TPI fusion enzymes will decrease as the change in [S] is no longer proportional to rate. However the mutations did not cause a rate increase of 22 fold, there was only a maximum of a 5.6 fold difference between the wtTPI-GAPDH and TPI-GAPDH H96A E165A (Table 6).

The change in apparent rate of GAPDH in the mutant TPI fusion proteins was seen because there was reduced or no TPI activity to perform the reverse reaction of GAP back into DHAP. This meant the substrate availability of GAP for GAPDH in the assay mix was increased and the apparent [S] was different between the wtTPI-GAPDH assay and the mutant TPI fusion protein assays due to the differences in the TPI activity. This led to an apparent increase in GAPDH rate, therefore decrease in GAPDH K_M for the mutant TPI fusion proteins as seen in Table 5 and 6. Both the E165A and H96A mutations effect the key catalytic sites on the enzyme (Figure 23); E165 (E167 in the sequence alignment – Appendix 2) is involved in the initiation of the reaction by removal of a proton from the substrate, as well as completing the reaction by transferring the proton back to the isomerised product, whilst H96 (H95 in the sequence alignment – Appendix 2) is involved in proton mediation (Samanta et al., 2011; Wierenga et al., 2010).

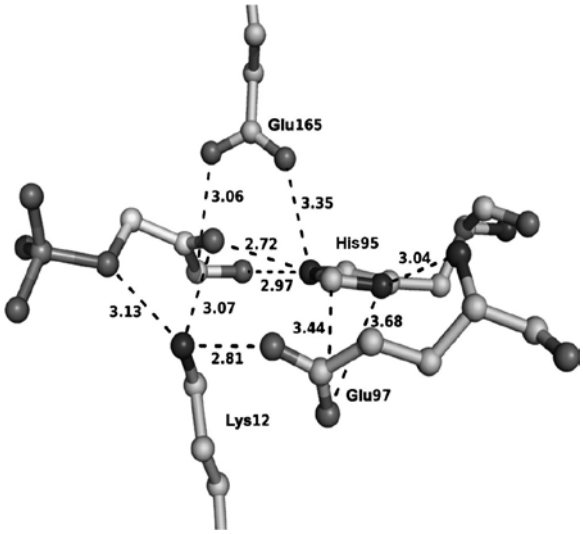


Figure 23: Active site residues of *Plasmodium falciparium* TPI. Showing the active residues for this organism at E165 and H95.
From Samanta et al, 2011. Revisiting the mechanism of the triosephosphate isomerase reaction: the role of the fully conserved glutamic acid 97 residue

It was surprising that, whilst showing very low rates and high K_Ms, the mutant TPI fusion proteins (especially TPI-GAPDH H96A E165A) had any rate at all with DHAP with such important residues mutated (Table 8, Figure 20). It was also interesting to see that whilst there was a significant increase in rate and decrease in K_M with GAP as a substrate, it was not proportional to the negative effects when DHAP was used as a substrate. For example there is a 76-fold difference between the activity of wtTPI-GAPDH and the less functionally active TPI-GAPDH H96A E165A when DHAP is used as a substrate (Table 8) but only a 5.6-fold difference between the two enzymes in TPI-GAPDH H96A E165A favour when GAP is used as a substrate (Table 6). This demonstrates that the TPI mutations are causing a change in the apparent K_M of GAPDH (apparent K_M because GAPDH is not working any faster, there is just less TPI activity limiting the substrate availability of GAP for GAPDH) but are not leaving TPI completely dysfunctional.

As seen in Table 7, pre-equilibrating the mutant TPI fusion enzymes with TPI reduces the rate compared to the control assays. The highest difference being a 13 times decrease from TPI-GAPDH E165A, rather than TPI-GAPDH H96A E165A

(which has the lowest K_M, so should have been the most effective mutation, however TPI-GAPDH E165A had a lower average rate with the pre-incubated TPI assay, this has been reflected in the normalised data), again suggesting the mutation is not a complete knockout to stop activity of the enzyme, even with the double mutation.

Pre-equilibrating the mutant TPI fusion proteins with the same concentration of TPI was effectively replacing the mutated TPI, however, as seen in Figure 20, there is still some activity within the mutated TPI with DHAP as a substrate. This could potentially explain the lower rates than wtTPI-GAPDH (0.16 mM mM^{-1} s^{-1}) when pre-equilibrated with TPI if there is still some latent TPI activity within the mutation; therefore increasing the concentration of TPI within the mix compared to wtTPI-GAPDH (Figure 19).

There are some issues with the data collected, especially when looking at Table 7. wtTPI-GAPDH and the mutant TPI fusion enzymes in the pre-equilibrated with TPI assays should have the same rate, since there is the same pre-equilibrated [S], the same [E] and an unchanged GAPDH (therefore same V_{max}). The differences are comparatively small (less than a factor of 2). This either suggests a potential difference in active [E] between the enzymes or larger than expected experimental variations.

A further source of potential error is in the low rates of the DHAP as a substrate assays. The very low rates collected (i.e. all below 0.02 mM mM^{-1} s^{-1}) means that small variations in rate across [S] can have a large apparent effect on the calculated V_{max} and K_M. The assays done with very low rates are therefore not as reliable as those with higher rates.

However, the most significant potential source of error in the calculations is in the GAP substrate used. In solution GAP has a tendency to form a geminal diol as well as the free aldehyde form at a molar ratio of 29:1 respectively. GAPDH cannot convert this geminal diol form into 1,3 BPG, so effectively the [S] of GAP in solution

is reduced by 29 fold. Presuming that TPI does not utilise the geminal diol form in the same way that GAPDH does not, all assays that use GAP as a substrate could potentially be out by 29 fold. TPI must not be able to utilise the geminal diol form of GAP as the pre-equilibrated with TPI assay rates would increase as TPI would be able to convert the geminal diol form to aldehyde form, showing a net increase in rate, however this is not seen (Trentham et al., 1969).

4.3 Separating *B. hominis* GAPDH from the TPI moiety causes a reduction in function in GAPDH activity. It decreases in apparent V_{max} and increases the apparent K_M.

As seen throughout the Chapter 3, the separated GAPDH His C and His N are consistently inferior enzymes to the wtTPI-GAPDH. Figure 16 shows a comparison between the rates of these three enzymes, with GAPDH His N having the lowest initial rate of 0.04 mM mM^{-1} s^{-1} and wtTPI-GAPDH an initial rate of 0.35 mM mM^{-1} s^{-1}. The difference in K_M between GAPDH His C and wtTPI-GAPDH is 1.79-fold, however the Vmax is 2.11-fold less than wtTPI-GAPDH (Table 5). GAPDH His C is a slower enzyme when fully saturated than wtTPI-GAPDH, this shows just how badly effected by the separation from TPI that GAPDH is. At its maximum, the activity of separated GAPDH could be expected to increase by 22 fold without the effects of TPI converting GAP to DHAP, therefore increasing the available [S] for GAPDH. Since the rate and V_{max} of GAPDH His C drops this supports the hypothesis that protists have kept TPI-GAPDH fused together for a metabolic reason. It is believed that the protists have kept the enzyme this way as an evolutionary throw back on an ancestral *alpha proteobacterial* gene. Other organisms are presumed to have two separate enzymes from the Archael host. This along with the mitochondrial targeting pre-sequence in these protists, suggests a mitochondrial origin for glycolysis in these organisms (Liaud et al., 2000; Nakayama et al., 2012). Once split from TPI, GAPDH does not function well on its own. The drop in rate when GAPDH His C is put in a TPI pre-equilibrated assay mix is almost 22 fold (18 fold), showing the separated TPI enzyme is working essentially as expected. This also shows that

at 0.25mM of GAP, GAPDH His C is the furthest away from K_M compared to the other enzymes as it has the highest difference, confirming its high K_M value (Table 7).

There is a marked difference between GAPDH His C and His N. GAPDH His N is poorer of the two enzymes, being 8.8 times less functional than wtTPI-GAPDH compared to GAPDH His C which is only looses 3.8-fold of its activity at 0.25mM GAP (Table 7). This is most likely due to the position of the His tag. The N terminus of GAPDH attaches to the C terminus of TPI (Figure 24). The placement of the His tag at the N terminus of GAPDH may mean that any interaction that is required between TPI and GAPDH for GAPDH to function is be disrupted. It is possible that TPI causes some kind of structural or folding change by being able to interact with the N terminus of GAPDH that allows GAPDH to function more efficiently than when it is separated.

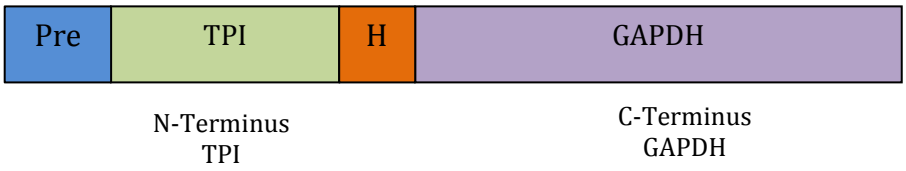


Figure 24: Diagram of the TPI-GAPDH protein. The C terminus of the TPI attaches to the N terminus of GAPDH via a hinge region (orange). The mitochondrial targeting pre-sequence is also shown (blue) (Liaud et al. 2000).

4.4 The TPI moiety is in competition with the GAPDH portion of TPI-GAPDH for GAP

The general assumptions in the literature and used in modelling the glycolytic flux expects TPI to have a fast enough turnover rate to create an 22:1 DHAP to GAP equilibrium almost immediately**.** There is no mention of GAPDH when TPI kinetics are discussed, nor any mention of TPI's effect on GAPDH in glycolytic kinetic studies. Glycolytic flux studies also show GAPDH to be an important regulator in

glycolysis (Albert et al., 2005; Stanford et al., 2013; Wierenga et al., 2010). This makes the results seen in 3.2.4 interesting and novel. As can be seen in Figure 19; pre-incubating the assay mix with TPI does not slow down the rate of wtTPI-GAPDH by 22-fold, as it almost does with GAPDH His C, it is only half as slow. This could be because wtTPI-GAPDH is closer to K_M and V_{max} in the M-M curve at 0.25mM GAP (Figure 13) than GAPDH His C, or potentially because TPI is competing with GAPDH for GAP. If the GAPDH portion is competing with the TPI portion for GAP, it could artificially make the rate appear higher in the control wtTPI-GAPDH assay compared to the pre-equilibrated assay than it would be if the TPI and GAPDH acted in the expected way (i.e. TPI creating an immediate equilibrium). This is also demonstrated in Figure 19 and Table 7 when comparing the pre-equilibrated wtTPI-GAPDH assay to pre-mixing the enzyme with extra TPI. Adding more TPI means there is a higher concentration of TPI; referring back to 3.2.2, increasing the [E] will increase the rate of reaction if the [S] remains the same. In the case of the extra TPI assays it decreases the rate at which GAPDH produces product by 1.8-fold because the extra TPI is competing with GAPDH for GAP and converting it to DHAP, therefore removing available substrate for GAPDH. This shows that the decrease in rates observed is due to the decrease in available [S] for GAPDH rather than a change in V_{max} or K_M. The addition of extra TPI and the pre-equilibration assays for the wtTPI-GAPDH and the mutant TPI fusion enzymes should have the same rates if TPI was equilibrating the substrate instantly. The rate decreases with the addition of more TPI (therefore an increase of TPI concentration), showing that TPI is working competitively for GAP against GAPDH.

Making GAPDH an inferior enzyme by splitting it apart from TPI, appears to make TPI a better enzyme, as seen in the pre-equilibration assays (Figure 19, Table 7). Pre-equilibrating wtTPI-GAPDH with TPI gives a 2-fold decrease in rate compared to the standard assay, whilst pre-equilibrating GAPDH His C with TPI shows a 18-fold decrease in rate compared to the standard assay. However, this does not mean TPI as an enzyme has improved. TPI only appears to have a faster rate because it does not have to compete with a fully functional GAPDH molecule for GAP in the separate GAPDH enzymes unlike it does when it is attached to GAPDH in the fusion

enzyme. This also applies to GAPDH when attached to a mutant TPI, the actual turnover rate of the enzyme does not change, and it is no longer competing with TPI for GAP so it has a higher [S] available to it, therefore has a lower apparent K_M.

A further result supporting GAPDH competing with TPI for GAP is that no burst period is seen in the enzymatic assays. Figures 14, 15 and 16 show that the reactions all occurred in a steady progression, producing product in a smooth and continuous manner from their start. The lack of an initial burst period of rate which then slowed down to a controlled rate, as would be expected if TPI was creating a substrate equilibrium almost, but not quite, immediately. There is not a pre-equilibrium as shown by the results of the assays in Table 7.

It is possible the reason GAPDH competing with TPI for GAP has not been discussed in the literature is because this is entirely novel to TPI-GAPDH, the fact that they are joined together means they have to outcompete each other. Also, cells have control over the ratio of enzymes; therefore it is possible the cells will express higher levels of TPI to GAPDH to ensure the 22:1 substrate ratio is maintained. With the enzymes being joined together it is impossible to control the ratio of enzyme expression.

4.5 Preliminary results suggest the potential for a functional dimeric GAPDH.

Some preliminary results were obtained from the size exclusion chromatography of *B. hominis* TPI-GAPDH, TPI and GAPDH (Figure22, Table 9). Monomeric (68kDa), dimeric (141kDa) and tetrameric (231kDa) forms of TPI-GAPDH were observed, suggesting unfavourable conditions for the enzyme, causing it to either form aggregates or fall apart. The dimeric form of TPI-GAPDH would be indicative of a dimer of TPI and GAPDH dimers structure, however there still is a small presence of tetrameric TPI-GAPDH. With the large peak of monomeric TPI-GAPDH it cannot be ruled out that much of tetrameric structure has fallen apart as opposed to the

dimers of TPI-GAPDH aggregating together. Though there is also apparently a small peak of potentially tetrameric TPI that eluted earlier than dimeric TPI. As there is no known tetrameric TPI, aggregate forming is possible.

However, the interesting result is with the GAPDH chromatograph, as this does not show a tetrameric GAPDH form, only a potential dimeric peak giving a mwt of 64kDa supporting the hypothesis that TPI-GAPDH is a dimer of dimers. This is also a key finding, as glycolytically active dimeric GAPDH has not been discovered before, only dimeric GAPDH that does other functions as mentioned in 1.2.2 (Tristan et al., 2011). Stabilising the protein samples to not fall apart or form aggregates is a key step in any future work into the structural determination of TPI-GAPDH.

4.6 Future work

As there appears to be inconsistencies in the data (e.g. Figure 19: TPI-GAPDH E165A having a higher rate than TPI-GAPDH H96A E165A), reproducing the work and repeating the assays a number of times would determine whether the values are true or erroneous. Also, there are key places in which more data over a greater range of [S] would have been beneficial. For example, during the pre-incubation with TPI assays data was only collected for 0.25mM of GAP rather than over a range of GAP concentrations (Table 7). Therefore the V_{max} and K_M values could not be calculated and a curve not fitted. This data would have told us how much pre-equilibrating with TPI affects the apparent K_M of an enzyme compared to the normal enzyme assays. This would give a better understanding of how much TPI competes with GAPDH.

Another issue is the range of [S] used for the assays was not representative enough; higher [S] would give rates that extended past the linear region of the M-M curve (Figure 12) towards the V_{max} and K_M. Table 5 and 6 show that the maximum [S] used (0.5mM) did not reach the K_M of the enzymes assayed (excluding TPI-

GAPDH C404A and GAPDH His N). Higher [S] would allow for more realistic K_Ms and curve fittings. The small [S] in the linear region can lead to more errors within the data, as a small error can have a large effect on the simulations that estimate V_{max} and K_M.

A further potential source of error would be in the number of active enzymes present. All of the assays contained the same [E] based on absorbance read at 280nm by spectrophotometry. This measures the absorbance of aromatic rings on amino acids in proteins and does not differentiate between the enzyme of interest and other proteins that may have made it through purification. A potential way round this would be to re-purify the protein fractions by the Protino affinity chromatography column. This may cause a loss of yield but should make for a purer protein sample. Another source of [E] difference could be in enzyme degradation, leading to a loss of active sites, this could be due to mishandling of the proteins outside of their favoured conditions. Checking active enzyme activity can be done by comparing the activity against the activity of a well-characterised GAPDH enzyme from another organism under the same conditions.

Further work would include repeating the protein production from the clones and all of the enzymatic assays again to show that the results gained are a reliable and a realistic representation of the characterisation of TPI-GAPDH and its mutations. Metabolic flux analysis and stop-flow enzymatic assays may give more understanding of the competitive relationship between TPI and GAPDH. Further size exclusion chromatography and x-ray crystallography will give a fuller understanding to the structure of the TPI-GAPDH enzyme; ultimately deciding whether it is a GAPDH tetramer with 2 TPI dimers attached to each end or a dimer of GAPDH and TPI dimers. If the structure is a dimer of dimers as suggested by the preliminary data discussed in 4.5, then this is a novel discovery, as no glycolytically active GAPDH dimers have been discovered yet. Understanding the structure of TPI-GAPDH will allow us to expand on its potential as a drug target and investigations into chemically breaking the hinge region between TPI and GAPDH may be noteworthy.

4.7 Conclusion

A selection of protist parasites, part of the *Stramenopiles*, contains a glycolytically active mitochondrial targeted TPI-GAPDH fusion enzyme. Mutations in the TPI portion of the fusion enzyme allows GAPDH to appear to have a higher rate and lower K_M, when in fact it just has a higher [S] available to it, as the mutated TPI cannot effectively convert GAP to DHAP. The TPI portion of the enzyme appears to compete with GAPDH for GAP, rather than immediately forming an equilibrium as expected. This is potentially due to the two enzymes being linked and the cells containing them not being able to adjust the ratio of TPI to GAPDH. However, this finding is still novel, as there has been no mention of this phenomenon in the literature.

TPI-GAPDH can only function effectively when together, as separating the two moieties dramatically decreases the function of the GAPDH portion of the enzyme, however it appears to have little effect on the functioning of TPI. Considering GAPDH without the influence of TPI could potentially be a maximum of 22 times faster at lower [S], the 1.79-fold increase in K_M and 2.11-fold decrease in V_{max} compared to TPI-GAPDH is significant. It appears that TPI has some effect on the structure of the GAPDH moiety catalytic site, as the position of the His tag used for protein collection had an impact on the rate of the enzyme. Mutating the TPI portion of the TPI-GAPDH fusion protein causes the apparent K_M of GAPDH to decrease, this coupled with the fact GAPDH does not function as well when split apart from TPI suggests the fusion protein remained fused due to structural reasons rather than kinetic reasons.

As treating protist parasites is fraught with difficulties through side effects and resistance; innovative ways of combatting the diseases are needed. The novel TPI-GAPDH enzyme poses a potential target for therapeutics when considering targeting the link between the two enzymes. This would decrease the functionality

of GAPDH and may have a negative effect on glycolysis. This could potentially be combined with mutations in the TPI portion of the enzyme, in which studies have already shown the survival rates of other protists decreasing when their TPI is affected with the mutations used in this project (Samanta et al., 2011). Knocking out two key glycolytic enzymes may be key to combatting protist parasite disease.

References

Albert, M. A., Haanstra, J. R., Hannaert, V., Van Roy, J., Opperdoes, F. R., Bakker, B. M. and Michels, P. A. (2005). Experimental and in silico analyses of glycolytic flux control in bloodstream form trypanosoma brucei. *J. Biol. Chem.* **280,** 28306-28315.

Alvarez, A., Martinez-Cadena, G., Silva, M., Avila, E. and Saavedra, E. (2007). Entamoeba histolytica: ADP-ribosylation of secreted glyceraldehyde-3-phosphate dehydrogenase. *Exp. Parasitol.* **117,** 349-356.

Banerjee, M., Balaram, H. and Balaram, P. (2009). Structural effects of a dimer interface mutation on catalytic activity of triosephosphate isomerase. *FEBS Journal* **276,** 4169-4183.

Banner, D. W., Bloomer, A. C., Petsko, G. A., Phillips, D. C. and Wilson, I. A. (1976). Atomic coordinates for triose phosphate isomerase from chicken muscle. *Biochem. Biophys. Res. Commun.* **72,** 146-155.

Bell, C. A., Cory, M., Fairley, T. A., Hall, J. E. and Tidwell, R. R. (1991). Structure-activity relationships of pentamidine analogs against giardia lamblia and correlation of antigiardial activity with DNA-binding affinity. *Antimicrob. Agents Chemother.* **35,** 1099-1107.

Berg, J. M., Tymoczko, J. L. and Stryer, L. (2007). Glycolysis and cluconeogenesis. In *Biochemistry (6th Edition)*: WH Freeman.

Bero, J., Beaufay, C., Hérent, M., Quetin-Leclercq, J., Hannaert, V. and Michels, P. (2013). Antitrypanosomal compounds from the essential oil and extracts of keetia leucantha leaves with inhibitor activity on trypanosoma brucei glyceraldehyde-3-phosphate dehydrogenase. *Phytomedicine* **20,** 270-274.

Branlant, G. and Branlant, C. (1985). Nucleotide sequence of the escherichia coli gap gene. *European Journal of Biochemistry* **150,** 61-66.

Bringaud, F., Rivière, L. and Coustou, V. (2006). Energy metabolism of trypanosomatids: Adaptation to available carbon sources. *Mol. Biochem. Parasitol.* **149,** 1-9.

Cáceres, A., Michels, P. and Hannaert, V. (2010). Genetic validation of aldolase and glyceraldehyde-3-phosphate dehydrogenase as drug targets in trypanosoma brucei. *Mol. Biochem. Parasitol.* **169,** 50-54.

Cardoso, C. L., de Moraes, M. C., Guido, R. V. C., Oliva, G., Andricopulo, A. D., Wainer, I. W. and Cass, Q. B. (2008). The development of an immobilized enzyme reactor containing glyceraldehyde-3-phosphate dehydrogenase from trypanosoma cruzi: The effect of species' specific differences on the immobilization. *Analyst* **133,** 93-99.

Chuang, D., Hough, C. and Senatorov, V. V. (2005). Glyceraldehyde-3-phosphate dehydrogenase, apoptosis, and neurodegenerative diseases. *Annu. Rev. Pharmacol. Toxicol.* **45,** 269-290.

Clos, J. and Brandau, S. (1994). pJC20 and pJC40-two high-copy-number vectors for T7 RNA polymerase-dependent expression of recombinant genes in< i> escherichia coli. *Protein Expr. Purif.* **5,** 133-137.

Cooke, D. E. L., Cano, L. M., Raffaele, S., Bain, R. A., Cooke, L. R., Etherington, G. J., Deahl, K. L., Farrer, R. A., Gilroy, E. M., Goss, E. M. et al. (2012). Genome analyses of an aggressive and invasive lineage of the irish potato famine pathogen. *PLoS Pathogens* **8,** 1-14.

Cudmore, S. L., Delgaty, K. L., Hayward-McClelland, S., Petrin, D. P. and Garber, G. E. (2004). Treatment of infections caused by metronidazole-resistant trichomonas vaginalis. *Clin. Microbiol. Rev.* **17,** 783.

Darkin-Rattray, S., Gurnett, A. M., Myers, R. W., Dulski, P. M., Crumley, T. M., Allocco, J. J., Cannova, C., Meinke, P. T., Colletti, S. L., Bednarek, M. A. et al. (1996). Apicidin: A novel antiprotozoal agent that inhibits parasite histone deacetylase. *Proc. Natl. Acad. Sci. U. S. A.* **93,** 13143-13147.

Davidson, B., Sajgo, M., Noller, H. and Harris, J. I. (1967). Amino-acid sequence of glyceraldehyde 3-phosphate dehydrogenase from lobster muscle. *Nature* **216,** 1181-1185.

Der Giezen, M. V. (2011). Mitochondria and the rise of eukaryotes. *Bioscience* **61,** 594-601.

Embley, T. M. and Martin, W. (2006). Eukaryotic evolution, changes and challenges. *Nature* **440,** 623-630.

Fry, W. (2008). Phytophthora infestans: The plant (and R gene) destroyer. *Molecular plant pathology*.

Ginger, M. L. (2006). Niche metabolism in parasitic protozoa. *Philos Trans R Soc Lond B Biol Sci* **361,** 101-118.

Glaser, P. E., Han, X. and Gross, R. W. (2002). Tubulin is the endogenous inhibitor of the glyceraldehyde 3-phosphate dehydrogenase isoform that catalyzes membrane fusion: Implications for the coordinated regulation of glycolysis and membrane fusion. *Proc. Natl. Acad. Sci. U. S. A.* **99,** 14104-14109.

Haas, B. J., Kamoun, S., Zody, M. C., Jiang, R. H. Y., Handsaker, R. E., Cano, L. M., Grabherr, M., Kodira, C. D., Raffaele, S., Torto-Alalibo, T. et al. (2009). Genome sequence and analysis of the irish potato famine pathogen phytophthora infestans. *Nature* **461,** 393-398.

Hammarton, T. C., Mottram, J. C. and Doerig, C. (2003). The cell cycle of parasitic protozoa: Potential for chemotherapeutic exploitation. *Prog. Cell Cycle Res.* **5,** 91-101.

Haresh, K., Suresh, K., Khairul Anus, A. and Saminathan, S. (1999). Isolate resistance of blastocystis hominis to metronidazole. *Trop. Med. Int. Health* **4,** 274-277.

Harris, R. (2006). Carbohydrate metabolism I: Major metabolic pathways and their control. In *Textbook of Biochemistry with Clinical Correlations (6th Edition)* (ed. T. Devlin), pp. 587-595: Wiley-Liss.

Haverkort, A. J., Boonekamp, P. M., Hutten, R., Jacobsen, E., Lotz, L. A. P., Kessel, G. J. T., Visser, R. G. F. and van, d. V. (2008). Societal costs of late blight in potato and prospects of durable resistance through cisgenic modification. *Potato Res.*

Hirata, T., Hokama, A., Fujita, J., Kinjo, N., Kinjo, F., Nakamura, H., Yamane, N. (2007). Prevalence of blastocystis hominis and strongyloides stercoralis infection in okinawa, japan. *Parasitol. Res.* **101,** 1717-1719.

Jacobs, R. T., Nare, B. and Phillips, M. A. (2011). State of the art in african trypanosome drug discovery. *Curr Top Med Chem* **11,** 1255-1274.

Jacq, C., Alt-Morbe, J., Andre, B., Arnold, W., Bahr, A., Ballesta, J. P., Bargues, M., Baron, L., Becker, A., Biteau, N. et al. (1997). The nucleotide sequence of saccharomyces cerevisiae chromosome IV. *Nature* **387,** 75-78.

Jones, G. and Harris, J. I. (1972). Glyceraldehyde 3-phosphate dehydrogenase: Amino acid sequence of enzyme from baker's yeast. *FEBS Lett.* **22,** 185-189.

Kennedy, P. G. (2013). Clinical features, diagnosis, and treatment of human african trypanosomiasis (sleeping sickness). *The Lancet Neurology* **12,** 186-194.

Kim, J. H., Lee, S., Park, J. B., Lee, S. D., Kim, J. H., Ha, S. H., Hasumi, K., Endo, A., Suh, P. and Ryu, S. H. (2003). Hydrogen peroxide induces association between glyceraldehyde 3-phosphate dehydrogenase and phospholipase D2 to facilitate phospholipase D2 activation in PC12 cells. *J. Neurochem.* **85,** 1228-1236.

Knowles, J. R. and Albery, W. J. (1977). Perfection in enzyme catalysis: The energetics of triosephosphate isomerase. *Acc. Chem. Res.* **10,** 105-111.

Kuntz, D. A., Osowski, R., Schudok, M., Wierenga, R. K., Müller, K., Kessler, H. and Opperdoes, F. R. (1992). Inhibition of triosephosphate isomerase from trypanosoma brucei with cyclic hexapeptides. *Eur. J. Biochem.* **207,** 441-447.

Kunz, S., Johner, A., Oberholzer, M., Wentzinger, L., Seebeck, T., Beavo, J. A., Laxman, S., D'Angelo, M.A., Flawia, M. M., Francis, S. H. et al. (2006). Cyclic nucleotide specific phosphodiesterases of the kinetoplastida: A unified nomenclature. *Mol. Biochem. Parasitol.* **145,** 133-135.

Lalle, M. (2010). Giardiasis in the post genomic era: Treatment, drug resistance and novel therapeutic perspectives. *Infect Disord Drug Targets* **10,** 283-294.

Leepin, A., Hemphill, A., Stüdli, A., Brun, R., Stephens, C. E. and Boykin, D. W. (2008). Host cells participate in the in vitro effects of novel diamidine analogues against tachyzoites of the intracellular apicomplexan parasites neospora caninum and toxoplasma gondii. *Antimicrob. Agents Chemother.* **52,** 1999-2008.

Lefevre, F., Remy, M. H. and Masson, J. M. (1997). Alanine-stretch scanning mutagenesis: A simple and efficient method to probe protein structure and function. *Nucleic Acids Res.* **25,** 447-448.

Lesk, A. (2010). *Introduction to Protein Science. 2nd Edition.* New York: Oxford University Press.

Li, L.H., Lv, S., Zhang, L., Tong, X. M., Chen, S. H., Zhou, X. N., Zhang, X. P., Yoshikawa, H., Wu, Z., Steinmann, P. et al. (2007). Cross-sectional surveys and subtype classification of human blastocystis isolates from four epidemiological settings in china. *Parasitol. Res.* **102,** 83-90.

Liaud, M., Martin, W., Cerff, R., Lichtlé, C. and Apt, K. (2000). Compartment-specific isoforms of TPI and GAPDH are imported into diatom mitochondria as a fusion protein: Evidence in favor of a mitochondrial origin of the eukaryotic glycolytic pathway. *Mol. Biol. Evol.* **17,** 213-223.

Li, C. and Xin, S. (2009). Discovery, mechanisms of action and combination therapy of artemisinin. *Expert Review of Anti-Infective Therapy* **7,** 999-1013.

Löfmark, S., Edlund, C. and Nord, C. E. (2010). Metronidazole is still the drug of choice for treatment of anaerobic infections. *Clin. Infect. Dis.* **50 Suppl 1,** S16-S23.

Lu, H. S., Yuan, P. M. and Gracy, R. W. (1984). Primary structure of human triosephosphate isomerase. *Journal of Biological Chemistry* **259,** 11958-11968.

Madigan, M., Martinko, J., Dunlap, P. and Clark, D. (2009). Antimicrobial agents used *in vivo*. In *Brock Biology of Microorganisms. 12th Edition* (ed. L. Berriman and G. Carlson), pp. 791-799. San Fransisco: Pearson Benjamin Cummings.

Mallo, N., Lamas, J. and Leiro, J. M. (2013). Hydrogenosome metabolism is the key target for antiparasitic activity of resveratrol against trichomonas vaginalis. *Antimicrob. Agents Chemother.* **57,** 2476-2484.

Marbach, A. and Bettenbrock, K. (2012). < I> lac operon induction in< i> escherichia coli: Systematic comparison of IPTG and TMG induction and influence of the transacetylase LacA. *J. Biotechnol.* **157,** 82-88.

Michels, P., Herman, M., Hannaert, V. and Bringaud, F. (2006). Metabolic functions of glycosomes in trypanosomatids. *Biochimica et Biophysica Acta - Molecular Cell Research* **1763,** 1463-1477.

Moras, D., Olsen, K. W., Sabesan, M. N., Buehner, M., Ford, G. C. and Rossmann, M. G. (1975). Studies of asymmetry in the three-dimensional structure of lobster D-glyceraldehyde-3-phosphate dehydrogenase. *J. Biol. Chem.* **250,** 9137-9162.

Mukherjee, S., Dutta, D., Saha, B. and Das, A. K. (2010). Crystal structure of glyceraldehyde-3-phosphate dehydrogenase 1 from methicillin-resistant staphylococcus aureus MRSA252 provides novel insights into substrate binding and catalytic mechanism. *J. Mol. Biol.* **401,** 949-968.

Müller, J. and Hemphill, A. (2011). Drug target identification in intracellular and extracellular protozoan parasites. *Curr Top Med Chem* **11,** 2029-2038.

Müller, M., Mentel, M., van Hellemond, J.,J., Henze, K., Woehle, C., Gould, S. B., Yu, R., van, d. G., Tielens, A. G. M. and Martin, W. F. (2012). Biochemistry and evolution of anaerobic energy metabolism in eukaryotes. *Microbiol. Mol. Biol. Rev.* **76,** 444-495.

Nakajima, H., Amano, W., Kubo, T., Fukuhara, A., Ihara, H., Azuma, Y. T., Tajima, H., Inui, T., Sawa, A. and Takeuchi, T. (2009). Glyceraldehyde-3-

phosphate dehydrogenase aggregate formation participates in oxidative stress-induced cell death. *J. Biol. Chem.* **284,** 34331-34341.

Nakayama, T., Ishida, K. and Archibald, J. M. (2012). Broad distribution of TPI-GAPDH fusion proteins among eukaryotes: Evidence for glycolytic reactions in the mitochondrion? *PLoS ONE* **7,** 1-12.

Nishimura, A., Morita, M., Nishimura, Y. and Sugino, Y. (1990). A rapid and highly efficient method for preparation of competent escherichia coli cells. *Nucleic Acids Res.* **18,** 6169.

Nowak, K., Kuczek, M., Ostropolska, L., Malarska, A., Wolny, M. and Baranowski, T. (1975). The covalent structure of glyceraldehyde-phosphate dehydrogenase from human muscles. isolation and amino acid sequences of peptides from tryptic digest. *Hoppe Seylers Z. Physiol. Chem.* **356,** 1181-1183.

Oesper, P. (1954). The mechanism of action of glyceraldehyde-3-phosphate dehydrogenase. *J. Biol. Chem.* **207,** 421-429.

Omar, A. (2006). The epidemiology of *Blastocystis hominis* in the united states. *Research Journal of Parasitology* **1,** 1-10.

Orosz, F., Oláh, J. and Ovádi, J. (2009). Triosephosphate isomerase deficiency: New insights into an enigmatic disease. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* **1792,** 1168-1174.

Pegelow, K., Gross, R., Pietrzik, K., Lukito, W., Richards, A. L. and Fryauff, D. J. (1997). Parasitological and nutritional situation of school children in the sukaraja district, west java, indonesia.. *Southeast Asian J Trop Med Public Health.* **28,** 173-190.

Pichersky, E., Gottlieb, L. and Hess, J. (1984). Nucleotide sequence of the triose phosphate isomerase gene of escherichia coli. *Molecular and General Genetics MGG* **195,** 314-320.

Rajamanikam, A. and Govind, S. K. (2013). Amoebic forms of blastocystis spp. - evidence for a pathogenic role. *Parasites and Vectors* **6,**.

Rayan, P., McDonnell, P. A. and Stenzel, D. (2005). The effects of saturated fatty acids on giardia duodenalis trophozoites in vitro. *Parasitol. Res.* **97,** 191-200.

Roger, A. J. and Hug, L. A. (2006). The origin and diversification of eukaryotes: Problems with molecular phylogenetics and molecular clock estimation. *Philosophical Transactions: Biological Sciences*, 1039.

Ronai, Z. (1993). Glycolytic enzymes as DNA binding proteins. *Int. J. Biochem.* **25,** 1073-1076.

Salomon, C. J. (2012). First century of chagas' disease: An overview on novel approaches to nifurtimox and benzonidazole delivery systems. *J. Pharm. Sci.* **101,** 888-894.

Samanta, M., Murthy, M., Balaram, H. and Balaram, P. (2011). Revisiting the mechanism of the triosephosphate isomerase reaction: The role of the fully conserved glutamic acid 97 residue. *ChemBioChem* **12,** 1886-1896.

Sarna, J. R., Furtado, S. and Brownell, A. K. (2013). Neurologic complications of metronidazole. *Can. J. Neurol. Sci.* **40,** 768-776.

Sirover, M. A. (1997). Role of the glycolytic protein, glyceraldehyde-3-phosphate dehydrogenase, in normal cell function and in cell pathology. *J. Cell. Biochem.* **66,** 133.

Stanford, N. J., Lubitz, T., Smallbone, K., Klipp, E., Mendes, P. and Liebermeister, W. (2013). Systematic construction of kinetic models from genome-scale metabolic networks. *PloS one* **8,** e79195.

Strange, R. N. and Scott, P. R. (2005). Plant disease: A threat to global food security. *Annu. Rev. Phytopathol.* **43,** 83-116.

Tamura, K., Battistuzzi, F.U., Billing-Ross, P., Murillo, O., Filipski, A. and Kumar, S. (2012). Estimating divergence times in large molecular phylogenies. *Proc. Natl. Acad. Sci. U. S. A.* **109,** 19333-19338.

Tan, K. S. W. (2008). New insights on classification, identification, and clinical relevance of blastocystis spp. *Clin. Microbiol. Rev.* **21,** 639-665.

Tarze, A., Deniaud, A., Le Bras, M., Maillier, E., Molle, D., Larochette, N., Zamzami, N., Jan, G., Kroemer, G. and Brenner, C. (2007). GAPDH, a novel

regulator of the pro-apoptotic mitochondrial membrane permeabilization. *Oncogene* **26,** 2606-2620.

Tielens, A. G. M., Grinsven, K. W. A., Henze, K., Hellemond, J. J. and Martin, W. (2010). Acetate formation in the energy metabolism of parasitic helminths and protists. *Int. J. Parasitol.* **40,** 387-397.

Tisdale, E. J., Azizi, F. and Artalejo, C. R. (2009). Rab2 utilizes glyceraldehyde-3-phosphate dehydrogenase and protein kinase C{iota} to associate with microtubules and to recruit dynein. *J. Biol. Chem.* **284,** 5876-5884.

Trentham, D., McMurray, C. and Pogson, C. (1969). The active chemical state of D-glyceraldehyde 3-phosphate in its reactions with D-glyceraldehyde 3-phosphate dehydrogenase, aldolase and triose phosphate isomerase. *Biochem. J.* **114,** 19-24.

Trentham, D. R. (1968). Aspects of the chemistry of D-glyceraldehyde 3-phosphate dehydrogenase. *Biochem. J.* **109,** 603-612.

Tristan, C., Shahani, N., Sedlak, T. W. and Sawa, A. (2011). The diverse functions of GAPDH: Views from different subcellular compartments. *Cell. Signal.*, 317.

Wierenga, R. K., Kapetaniou, E. G. and Venkatesan, R. (2010). Triosephosphate isomerase: A highly evolved biocatalyst. *Cell Mol. Life Sci.* **67,** 3961-3982.

Winter, R. W., Kelly, J. X., Smilkstein, M. J., Dodean, R., Hinrichs, D. and Riscoe, M. K. (2008). Antimalarial quinolones: Synthesis, potency, and mechanistic studies. *Exp. Parasitol.* **118,** 487-497.

Wiser, M. F. (2011). *Protozoa and Human Disease.*: Garland Science, Taylor and Francis Group, LLC.

Yakoob, J., Jafri, W., Jafri, N., Islam, M. and Asim Beg, M. (2004). **In vitro susceptibility of blastocystis hominis isolated from patients with irritable bowel syndrome.**. *Br J Biomed Sci.* **61,** 75-77.

Zhang, Z., Sugio, S., Komives, E. A., Liu, K. D., Knowles, J. R., Petsko, G. A. and Ringe, D. (1994). Crystal structure of recombinant chicken triosephosphate isomerase-phosphoglycolohydroxamate complex at 1.8-. ANG. resolution. *Biochemistry (N. Y.)* **33,** 2830-2837.

Appendices

Appendix 1: GAPDH Primary Sequence Alignment

```
                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                                    0         10         20         30
GAPDH [Lob      ---------- ---------- ---------X SKIGINGFGR IGRLVLRAAL SCG-AQVVAV
TPI_GAPDH       IDGFLVGGAS LKPEFVNIIN ANSTVQNAGP VSVGINGFGR IGRLVFRASQ PNPLVNVVAI
GAPDH [B.       ~~~~~~~~MG SSHHHHHHSS GLVPRGSHMP VSVGINGFGR IGRLVFRASQ PNPLVNVVAI
GAPDH [Hom      ---------- ---------- -------MGK VKVGVNGFGR IGRLVTRAAF NSGKVDIVAI
GAPDH [Sac      ---------- ---------- ---------M VRVAINGFGR IGRLVMRIAL SRPNVEVVAL
GAPDH [Esc      ---------- ---------- --------MA IKIGINGFGR IGRLVLRVAL GRKDIEVVAV

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                          40         50         60         70         80         90
GAPDH [Lob      NDPFIALEYM VYMFKYDSTH GVFKGEVKME DGALVVDGKK ITVFNEMKPE NIPWSKAGAE
TPI_GAPDH       NDPFITPDYM EYMIMHDSTH GPFQGTVKAE KDAIIVNGRR IVVSNEMDPK KIQWGAAGAE
GAPDH [B.       NDPFITPDYM EYMIMHDSTH GPFQGTVKAE KDAIIVNGRR IVVSNEMDPK KIQWGAAGAE
GAPDH [Hom      NDPFIDLNYM VYMFQYDSTH GKFHGTVKAE NGKLVINGNP ITIFQERDPS KIKWGDAGAE
GAPDH [Sac      NDPFITNDYA AYMFKYDSTH GRYAGEVSHD DKHIIVDGKK IATYQERDPA NLPWGSSNVD
GAPDH [Esc      NDPFIAPDYA AYMFKYDSTH GRYKGEVTAS GDDLVIDGHK IKVFQERDPA NIPWGKSGVD

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                         100        110        120        130        140        150
GAPDH [Lob      YIVESTGVFT TIEKASAHFK GGAKKVVISA PSADAPMFVC GVNLEKYSKD MTVVSNASCT
TPI_GAPDH       YIVESTGVFT AKDKAAQHLE GGAKKVVISA PSKDAPMFVM GVNNTTYTKD LKVVSNASCT
GAPDH [B.       YIVESTGVFT AKDKAAQHLE GGAKKVVISA PSKDAPMFVM GVNNTTYTKD LKVVSNASCT
GAPDH [Hom      YVVESTGVFT TMEKAGAHLQ GGAKRVIISA PSADAPMFVM GVNHEKYDNS LKIISNASCT
GAPDH [Sac      IAIDSTGVFK ELDTAQKHID AGAKKVVITA PSSTAPMFVM GVNEEKYTSD LKIVSNASCT
GAPDH [Esc      YVIESTGVFT KLEGAQKHID AGAKKVIITA PSADAPMFVV GVNEDKYTPD LKIISNASCT

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                         160        170        180        190        200        210
GAPDH [Lob      TNCLAPVAKV LHENFEIVEG LMTTVHAVTA TQKTVDGPSA KDWRGGRGAA QNIIPSSTGA
TPI_GAPDH       TNCLAPLAKI VNDKFGLKEG LMTTVHSVTS TQKVLDGPSK KDWRGGRSAC YNIIPSSTGA
GAPDH [B.       TNCLAPLAKI VNDKFGLKEG LMTTVHSVTS TQKVLDGPSK KDWRGGRSAC YNIIPSSTGA
GAPDH [Hom      TNCLAPLAKV IHDNFGIVEG LMTTVHAITA TQKTVDGPSG KLWRDGRGAL QNIIPASTGA
GAPDH [Sac      TNCLAPLAKV INDAFGIEEG LMTTVHSLTA TQKTVDGPSH KDWRGGRTAS GNIIPSSTGA
GAPDH [Esc      TNCLAPLAKV VNDTFGIEEG LMTTVHSITA TQKTVDGPSH KDWRGGRTAS GNIIPSSTGA

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                         220        230        240        250        260        270
GAPDH [Lob      AKAVGKVIPE LDGKLTGMAF RVPTPDVSVV DLTVRLGKEC SYDDIKAAMK TASEGPLQGF
TPI_GAPDH       AKAVGKVIPE LNGKLTGMSF RVPTEDVSVV DLTCTLKKPA TYEQIKAAVK EASETYMKGI
GAPDH [B.       AKAVGKVIPE LNGKLTGMSF RVPTEDVSVV DLTCTLKKPA TYEQIKAAVK EASETYMKGI
GAPDH [Hom      AKAVGKVIPE LDGKLTGMAF RVPTANVSVV DLTCRLEKPA KYDDIKKVVK QASEGPLKGI
GAPDH [Sac      AKAVGKVLPE LQGKLTGMAF RVPTVDVSVV DLTVKLNKET TYDEIKKVVK AAAEGKLKGV
GAPDH [Esc      AKAVGKVIPE LNGKLTGMSL RVPTTDVSVV DLTVRLKKAA SYEEIAQAIK KASEGPLKGV
```

```
             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                     280        290        300        310        320        330
GAPDH [Lob   LGYTEDDVVS SDFIGDNRSS IFDAKAGIQL SKTFVKVVSW YDNEFGYSQR VIDLLKHMQK
TPI_GAPDH    MGYVDYDVVS RDLLTCPYSS VFDAKAGIAL NDTFVKLVSW YDNEWGYSNR MVDLIQYMAK
GAPDH [B.    MGYVDYDVVS RDLLTCPYSS VFDAKAGIAL NDTFVKLVSW YDNEWGYSNR MVDLIQYMAK
GAPDH [Hom   LGYTEHQVVS SDFNSDTHSS TFDAGAGIAL NDHFVKLISW YDNEFGYSNR VVDLMAHMAS
GAPDH [Sac   LGYTEDAVVS SDFLGGSHSS IFDASAGIQL SPKFVKLVSW YDNEYGYSTR VVDLVEHVAK
GAPDH [Esc   LGYTEDAVVS TDFLGSSYSS IFDEKAGILL SPTFVKLISW YDNEYGYSTR VVDLLEHVAK


             ....|

GAPDH [Lob   VDSA-
TPI_GAPDH    VDRS*
GAPDH [B.    VDRS*
GAPDH [Hom   KE---
GAPDH [Sac   A----
GAPDH [Esc   ASA—
```

Legend:

Active residues: Cysteine 149
Histidine 176

GAPDH [Lob: GAPDH Lobster (Davidson et al., 1967)
TPI_GAPDH: Wild type TPI-GAPDH fusion protein from *B. hominis*, continues previous to GAPDH sequence. Nucleotide sequence in Appendix 4
GAPDH [B: Separated GAPDH from *B. hominis*. Nucleotide sequence in Appendix 5
GAPDH [Hom: GAPDH *Homo sapiens* (Nowak et al., 1975)
GAPDH [Sac: GAPDH Saccharomyces cerevisiae (Jones and Harris, 1972)
GAPDH [Esc: GAPDH *Escherichia coli* (Branlant and Branlant, 1985)

Appendix 2: TPI Primary Sequence Alignment

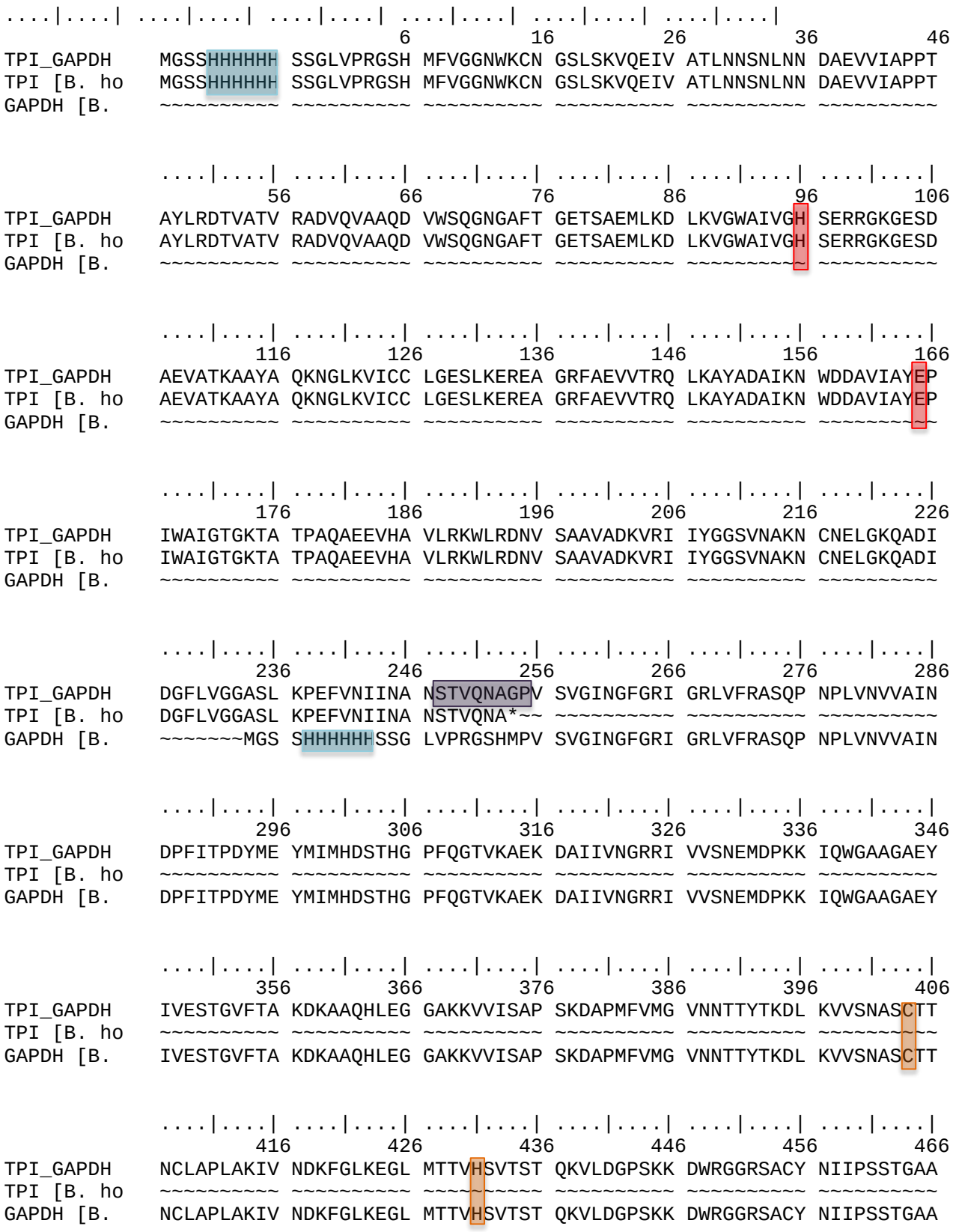
```
                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                                    0          10         20         30         40
TPI [G. ga   ---------- ------MAP- -RKFFVGGNW KMNGDKKSLG ELIHTLNGAK LSAD-TEVVC
TPI_GAPDH    ~~~MGSSHHH HHHSSGLVPR GSHMFVGGNW KCNGSLSKVQ EIVATLNNSN LNND~AEVVI
TPI [B. ho   ~~~MGSSHHH HHHSSGLVPR GSHMFVGGNW KCNGSLSKVQ EIVATLNNSN LNND~AEVVI
TPI [Homo    ---------- ------MAP- SRKFFVGGNW KMNGRKQSLG ELIGTLNAAK VPAD-TEVVC
TPI [Sacch   ---------- ------MAR- --TFFVGGNF KLNGSKQSIK EIVERLNTAS IPEN-VEVVI
TPI [Esche   ---------- ---------- MRHPLVMGNW KLNGSRHMVH ELVSNLRKEL AGVAGCAVAI

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                         50         60         70         80         90        100
TPI [G. ga   GAPSIYLDFA RQKLD-AKIG VAAQNCYKVP KGAFTGEISP AMIKDIGAAW VILGHSERRH
TPI_GAPDH    APPTAYLRDT VATVR~ADVQ VAAQDVWSQG NGAFTGETSA EMLKDLKVGW AIVGHSERRG
TPI [B. ho   APPTAYLRDT VATVR~ADVQ VAAQDVWSQG NGAFTGETSA EMLKDLKVGW AIVGHSERRG
TPI [Homo    APPTAYIDFA RQKLD-PKIA VAAQNCYKVT NGAFTGEISP GMIKDCGATW VVLGHSERRH
TPI [Sacch   CPPATYLDYS VSLVKKPQVT VGAQNAYLKA SGAFTGENSV DQIKDVGAKW VILGHSERRS
TPI [Esche   APPEMYIDMA KREAEGSHIM LGAQNVDLNL SGAFTGETSA AMLKDIGAQY IIIGHSERRT

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                        110        120        130        140        150        160
TPI [G. ga   VFGESDELIG QKVAHALAEG LGVIACIGEK LDEREAGITE KVVFEQTKAI AD--NVKDWS
TPI_GAPDH    K~GESDAEVA TKAAYAQKNG LKVICCLGES LKEREAGRFA EVVTRQLKAY AD~~AIKNWD
TPI [B. ho   K~GESDAEVA TKAAYAQKNG LKVICCLGES LKEREAGRFA EVVTRQLKAY AD~~AIKNWD
TPI [Homo    VFGESDELIG QKVAHALAEG LGVIACIGEK LDEREAGITE KVVFEQTKVI AD--NVKDWS
TPI [Sacch   YFHEDDKFIA DKTKFALGQG VGVILCIGET LEEKKAGKTL DVVERQLNAV LE--EVKDWT
TPI [Esche   YHKESDELIA KKFAVLKEQG LTPVLCIGET EAENEAGKTE EVCARQIDAV LKTQGAAAFE

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                        170        180        190        200        210        220
TPI [G. ga   KVVLAYEPVW AIGTGKTATP QQAQEVHEKL RGWLKSHVSD AVAQSTRIIY GGSVTGGNCK
TPI_GAPDH    DAVIAYEPIW AIGTGKTATP AQAEEVHAVL RKWLRDNVSA AVADKVRIIY GGSVNAKNCN
TPI [B. ho   DAVIAYEPIW AIGTGKTATP AQAEEVHAVL RKWLRDNVSA AVADKVRIIY GGSVNAKNCN
TPI [Homo    KVVLAYEPVW AIGTGKTATP QQAQEVHEKL RGWLKSNVSD AVAQSTRIIY GGSVTGATCK
TPI [Sacch   NVVVAYEPVW AIGTGLAATP EDAQDIHASI RKFLASKLGD KAASELRILY GGSANGSNAV
TPI [Esche   GAVIAYEPVW AIGTGKSATP AQAQAVHKFI RDYIAK-VDA NIAEQVIIQY GGSVNASNAA

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                        230        240        250        260        280        290
TPI [G. ga   ELASQHDVDG FLVGGASLKP -EFVDIINAK H--------- ---------- ----------
TPI_GAPDH    ELGKQADIDG FLVGGASLKP ~EFVNIINAN STVQNAGPVS VGINGFGRIG RLVFRASQPN
TPI [B. ho   ELGKQADIDG FLVGGASLKP ~EFVNIINAN STVQNA*~~~ ~~~~~~~~~~ ~~~~~~~~~~
TPI [Homo    ELASQPDVDG FLVGGASLKP -EFVDIINAK Q--------- ---------- ----------
TPI [Sacch   TFKDKADVDG FLVGGASLKP -EFVDIINSR KLRLI----- ---------- ----------
TPI [Esche   ELFAQPDIDG ALVGGASLKA DAFAVIVKAA EAAKQA---- ---------- ----------
```

Legend

Active residues: Histidine 95
Cysteine 167

TPI [G. ga: TPI *Gallus gallus* (Banner et al., 1976)
TPI_GAPDH: Wild type TPI-GAPDH fusion protein from *B. hominis*, sequence continues after TPI. Nucleotide sequence in Appendix 4
TPI [B. ho: TPI *B. hominis* Separated TPI from *B. hominis*. Nucleotide sequence in Appendix 6
TPI [Homo: TPI *Homo sapiens* (Lu et al., 1984)
TPI [Sacch: TPI *Saccharomyces cerevisiae* (Jacq et al., 1997)
TPI [Esche: TPI *Escherichia coli* (Pichersky et al., 1984)

Appendix 3: Wild Type TPI-GAPDH With Separated GAPDH and TPI Moieties Primary Sequence Alignment



```
                .....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                       476        486        496        506        516        526
TPI_GAPDH       KAVGKVIPEL NGKLTGMSFR VPTEDVSVVD LTCTLKKPAT YEQIKAAVKE ASETYMKGIM
TPI [B. ho      ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
GAPDH [B.       KAVGKVIPEL NGKLTGMSFR VPTEDVSVVD LTCTLKKPAT YEQIKAAVKE ASETYMKGIM

                .....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                       536        546        556        566        576        586
TPI_GAPDH       GYVDYDVVSR DLLTCPYSSV FDAKAGIALN DTFVKLVSWY DNEWGYSNRM VDLIQYMAKV
TPI [B. ho      ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
GAPDH [B.       GYVDYDVVSR DLLTCPYSSV FDAKAGIALN DTFVKLVSWY DNEWGYSNRM VDLIQYMAKV


                ....

TPI_GAPDH       DRS*
TPI [B. ho      ~~~~
GAPDH [B.       DRS*
```

Legend

Active residues: Histidine 96 (TPI)
Cysteine 165 (TPI)
Cysteine 404 (GAPDH)
Histidine 431 (GAPDH)

Residues to be mutated: Histidine 96 → Alanine (TPI)
Cysteine 165 → Alanine (TPI)
Cysteine 404 → Alanine (GAPDH)

TPI_GAPDH: Wild type TPI-GAPDH fusion protein from *B. hominis*, sequence continues after TPI. Nucleotide sequence in Appendix 4
TPI [B. ho: TPI *B. hominis* Separated TPI from *B. hominis*. Nucleotide sequence in Appendix 5
GAPDH [B: Separated GAPDH from *B. hominis*. Nucleotide sequence in Appendix 6

TPI active residues

GAPDH active residues

Hexa-his tags

Linker region

Appendix 4: wtTPI-GAPDH Sequence from Mark Van Der Giezen

ATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCA
TATGTTCGTCGGTGGCAATTGGAAGTGCAACGGTTCTCTCTCCAAGGTCCAGGAGATCG
TGGCTACTCTCAACAACTCCAACCTCAACAACGACGCTGAGGTCGTGATCGCCCCTCCC
ACCGCCTATCTGAGAGACACCGTGGCTACCGTGCGCGCTGATGTGCAGGTGGCTGCTCA
GGACGTCTGGTCCCAGGGCAACGGTGCCTTCACCGGCGAGACCTCTGCTGAAATGCTGA
AGGATCTCAAGGTCGGTTGGGCGATCGTGGGCCACTCCGAGCGCCGTGGCAAGGGCGAG
AGCGATGCCGAGGTTGCGACCAAGGCTGCTTACGCTCAGAAGAACGGTCTGAAGGTGAT
CTGCTGCCTGGGCGAGTCCCTGAAGGAGCGCGAGGCTGGCCGCTTCGCCGAGGTGGTGA
CCCGCCAGCTCAAGGCCTACGCCGACGCCATCAAGAACTGGGATGACGCCGTGATCGCC
TACGAGCCCATCTGGGCCATTGGCACGGGCAAGACCGCCACCCCCGCTCAGGCGGAGGA
GGTGCACGCCGTACTGCGCAAGTGGCTGCGCGACAATGTGAGCGCCGCCGTGGCGGACA
AGGTGCGCATCATCTACGGAGGCTCCGTGAACGCGAAGAACTGCAACGAGCTGGGTAAG
CAGGCCGACATCGACGGCTTCCTGGTGGGCGGCGCTTCGCTGAAGCCCGAGTTCGTGAA
CATCATCAACGCCAACTCCACGGTGCAGAACGCGGGCCCCGTGAGCGTGGGTATCAACG
GCTTCGGCCGCATTGGTCGTCTGGTGTTCCGCGCCAGCCAGCCGAACCCGCTGGTGAAC
GTGGTGGCGATCAACGACCCGTTCATCACGCCCGACTACATGGAGTACATGATCATGCA
CGACTCCACGCACGGTCCGTTCCAGGGCACCGTGAAGGCTGAGAAGGACGCGATCATCG
TGAACGGCCGCCGCATCGTGGTGTCCAACGAGATGGACCCGAAGAAGATCCAGTGGGGC
GCTGCGGGTGCGGAGTACATCGTGGAGTCCACGGGCGTGTTCACGGCGAAGGACAAGGC
CGCGCAGCACCTGGAGGGCGGCGCGAAGAAGGTGGTGATTTCCGCGCCGTCGAAGGACG
CTCCGATGTTCGTGATGGGCGTGAACAACACGACGTACACGAAGGACCTGAAGGTGGTG
TCGAACGCGTCCTGCACGACGAACTGCCTGGCGCCGCTGGCGAAGATCGTGAATGACAA
GTTCGGTCTGAAAGAGGGTCTGATGACGACGGTGCACTCGGTGACCTCCACGCAAAAGG
TGCTGGACGGTCCTTCCAAGAAGGACTGGAGAGGCGGCCGCTCTGCCTGCTACAACATC
ATCCCGTCCTCCACGGGCGCCGCCAAGGCCGTGGGCAAGGTCATTCCCGAGCTGAACGG
CAAGCTGACGGGCATGTCCTTCCGCGTGCCGACGGAGGACGTCTCCGTGGTGGATCTGA
CCTGCACGCTGAAGAAGCCCGCCACCTACGAGCAGATCAAGGCTGCGGTGAAGGAGGCT
TCGGAGACGTACATGAAGGGCATTATGGGATACGTGGATTACGACGTGGTGTCTCGCG
ATCTGCTGACCTGCCCGTACTCTTCTGTCTTCGACGCGAAGGCAGGAATTGCCCTGAAC
GACACGTTCGTGAAGCTGGTTTCTTGGTACGACAACGAGTGGGGCTACTCCAACCGTAT
GGTCGACCTCATCCAGTACATGGCGAAGGTGGATCGCTCTTAA

Appendix 5: TPI Sequence supplied by Mark Van Der Giezen

ATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCA
TATGTTCGTCGGTGGCAATTGGAAGTGCAACGGTTCTCTCTCCAAGGTCCAGGAGATCG
TGGCTACTCTCAACAACTCCAACCTCAACAACGACGCTGAGGTCGTGATCGCCCCTCCC
ACCGCCTATCTGAGAGACACCGTGGCTACCGTGCGCGCTGATGTGCAGGTGGCTGCTCA
GGACGTCTGGTCCCAGGGCAACGGTGCCTTCACCGGCGAGACCTCTGCTGAAATGCTGA
AGGATCTCAAGGTCGGTTGGGCGATCGTGGGCCACTCCGAGCGCCGTGGCAAGGGCGAG
AGCGATGCCGAGGTTGCGACCAAGGCTGCTTACGCTCAGAAGAACGGTCTGAAGGTGAT
CTGCTGCCTGGGCGAGTCCCTGAAGGAGCGCGAGGCTGGCCGCTTCGCCGAGGTGGTGA
CCCGCCAGCTCAAGGCCTACGCCGACGCCATCAAGAACTGGGATGACGCCGTGATCGCC
TACGAGCCCATCTGGGCCATTGGCACGGGCAAGACCGCCACCCCCGCTCAGGCGGAGGA
GGTGCACGCCGTACTGCGCAAGTGGCTGCGCGACAATGTGAGCGCCGCCGTGGCGGACA
AGGTGCGCATCATCTACGGAGGCTCCGTGAACGCGAAGAACTGCAACGAGCTGGGTAAG
CAGGCCGACATCGACGGCTTCCTGGTGGGCGGCGCTTCGCTGAAGCCCGAGTTCGTGAA
CATCATCAACGCCAACTCCACGGTGCAGAACGCGTAA

Appendix 6: GAPDH Sequence supplied by Mark Van Der Giezen

ATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCA
TATGCCCGTGAGCGTGGGTATCAACGGCTTCGGCCGCATTGGTCGTCTGGTGTTCCGCG
CCAGCCAGCCGAACCCGCTGGTGAACGTGGTGGCGATCAACGACCCGTTCATCACGCCC
GACTACATGGAGTACATGATCATGCACGACTCCACGCACGGTCCGTTCCAGGGCACCGT
GAAGGCTGAGAAGGACGCGATCATCGTGAACGGCCGCCGCATCGTGGTGTCCAACGAGA
TGGACCCGAAGAAGATCCAGTGGGGCGCTGCGGGTGCGGAGTACATCGTGGAGTCCACG
GGCGTGTTCACGGCGAAGGACAAGGCCGCGCAGCACCTGGAGGGCGGCGCGAAGAAGGT
GGTGATTTCCGCGCCGTCGAAGGACGCTCCGATGTTCGTGATGGGCGTGAACAACACGA
CGTACACGAAGGACCTGAAGGTGGTGTCGAACGCGTCCTGCACGACGAACTGCCTGGCG
CCGCTGGCGAAGATCGTGAATGACAAGTTCGGTCTGAAAGAGGGTCTGATGACGACGG
TGCACTCGGTGACCTCCACGCAAAAGGTGCTGGACGGTCCTTCCAAGAAGGACTGGAGA
GGCGGCCGCTCTGCCTGCTACAACATCATCCCGTCCTCCACGGGCGCCGCCAAGGCCGTG
GGCAAGGTCATTCCCGAGCTGAACGGCAAGCTGACGGGCATGTCCTTCCGCGTGCCGAC
GGAGGACGTCTCCGTGGTGGATCTGACCTGCACGCTGAAGAAGCCCGCCACCTACGAGC
AGATCAAGGCTGCGGTGAAGGAGGCTTCGGAGACGTACATGAAGGGCATTATGGGATA
CGTGGATTACGACGTGGTGTCTCGCGATCTGCTGACCTGCCCGTACTCTTCTGTCTTCG
ACGCGAAGGCAGGAATTGCCCTGAACGACACGTTCGTGAAGCTGGTTTCTTGGTACGAC
AACGAGTGGGGCTACTCCAACCGTATGGTCGACCTCATCCAGTACATGGCGAAGGTGGA
TCGCTCTTAA

Appendix 7: TPI-GAPDH C404A Sequence

ATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCA
TATGTTCGTCGGTGGCAATTGGAAGTGCAACGGTTCTCTCTCCAAGGTCCAGGAGATCG
TGGCTACTCTCAACAACTCCAACCTCAACAACGACGCTGAGGTCGTGATCGCCCCTCCC
ACCGCCTATCTGAGAGACACCGTGGCTACCGTGCGCGCTGATGTGCAGGTGGCTGCTCA
GGACGTCTGGTCCCAGGGCAACGGTGCCTTCACCGGCGAGACCTCTGCTGAAATGCTGA
AGGATCTCAAGGTCGGTTGGGCGATCGTGGGCCACTCCGAGCGCCGTGGCAAGGGCGAG
AGCGATGCCGAGGTTGCGACCAAGGCTGCTTACGCTCAGAAGAACGGTCTGAAGGTGAT
CTGCTGCCTGGGCGAGTCCCTGAAGGAGCGCGAGGCTGGCCGCTTCGCCGAGGTGGTGA
CCCGCCAGCTCAAGGCCTACGCCGACGCCATCAAGAACTGGGATGACGCCGTGATCGCC
TACGAGCCCATCTGGGCCATTGGCACGGGCAAGACCGCCACCCCCGCTCAGGCGGAGGA
GGTGCACGCCGTACTGCGCAAGTGGCTGCGCGACAATGTGAGCGCCGCCGTGGCGGACA
AGGTGCGCATCATCTACGGAGGCTCCGTGAACGCGAAGAACTGCAACGAGCTGGGTAAG
CAGGCCGACATCGACGGCTTCCTGGTGGGCGGCGCTTCGCTGAAGCCCGAGTTCGTGAA
CATCATCAACGCCAACTCCACGGTGCAGAACGCGGGCCCCGTGAGCGTGAGTATCAACG
GCTTCGGCCGCATTGGTCGTCTGGTGTTCCGCGCCAGCCAGCCGAACCCGCTGGTGAAC
GTGGTGGCGATCAACGACCCGTTCATCACGCCCGACTACATGGAGTACATGATCATGCA
CGACTCCACGCACGGTCCGTTCCAGGGCACCGTGAAGGCTGAGAAGGACGCGATCATCG
TGAACGGCCGCCGCATCGTGGTGTCCAACGAGATGGACCCGAAGAAGATCCAGTGGGGC
GCTGCGGGTGCGGAGTACATCGTGGAGTCCACGGGCGTGTTCACGGCGAAGGACAAGGC
CGCGCAGCACCTGGAGGGCGGCGCGAAGAAGGTGGTGATTTCCGCGCCGTCGAAGGACG
CTCCGATGTTCGTGATGGGCGTGAACAACACGACGTACACGAAGGACCTGAAGGTGGTG
TCGAACGCGTCCGCCACGACGAACTGCCTGGCGCCGCTGGCGAAGATCGTGAATGACAA
GTTCGGTCTGAAAGAGGGTCTGATGACGACGGTGCACTCGGTGACCTCCACGCAAAAGG
TGCTGGACGGTCCTTCCAAGAAGGACTGGAGAGGCGGCCGCTCTGCCTGCTACAACATC
ATCCCGTCCTCCACGGGCGCCGCCAAGGCCGTGGGCAAGGTCATTCCCGAGCTGAACGG
CAAGCTGACGGGCATGTCCTTCCGCGTGCCGACGGAGGACGTCTCCGTGGTGGATCTGA
CCTGCACGCTGAAGAAGCCCGCCACCTACGAGCAGATCAAGGCTGCGGTGAAGGAGGCT
TCGGAGACGTACATGAAGGGCATTATGGGATACGTGGATTACGACGTGGTGTCTCGCG
ATCTGCTGACCTGCCCGTACTCTTCTGTCTTCGACGCGAAGGCAGGAATTGCCCTGAAC
GACACGTTCGTGAAGCTGGTTTCTTGGTACGACAACGAGTGGGGCTACTCCAACCGTAT
GGTC

Appendix 8: TPI-GAPDH H96A Sequence

ATGGGCAGCAGCCATCATCATCATCTCTTATTATTG~GCCTGTTGCCGCGCGGCAGCCA
TATGTTCAGTGGTGGCATTTGGACGTGCAACGGTTCTCTCTCCAAGGTCCAGGAGATCC
TGGCTACTCTCAACAACTCCTACCTCAACAACGACGCTGAGGTCGTGATCGCCCCTCCC
ACCGCCTATCTGAGAGACACCGTGGCTACCGTGCGCGCTGATGTGTATGTGGCTGCTCA
GGACGTCTGGTCCCATGGCAACGGTGCCTTCACCGGCGATACCTCTGCTGAAATGCTGA
AGGATCTCAAGGTCGGTTGGGCGATCGTGGGCGCCTCCGAGCGCCGTGGCAAGGGCGAG
AGCGATGCCGATGTTGCGTTCATTTGCTGCTTA

Sequence provided by K. Favas.

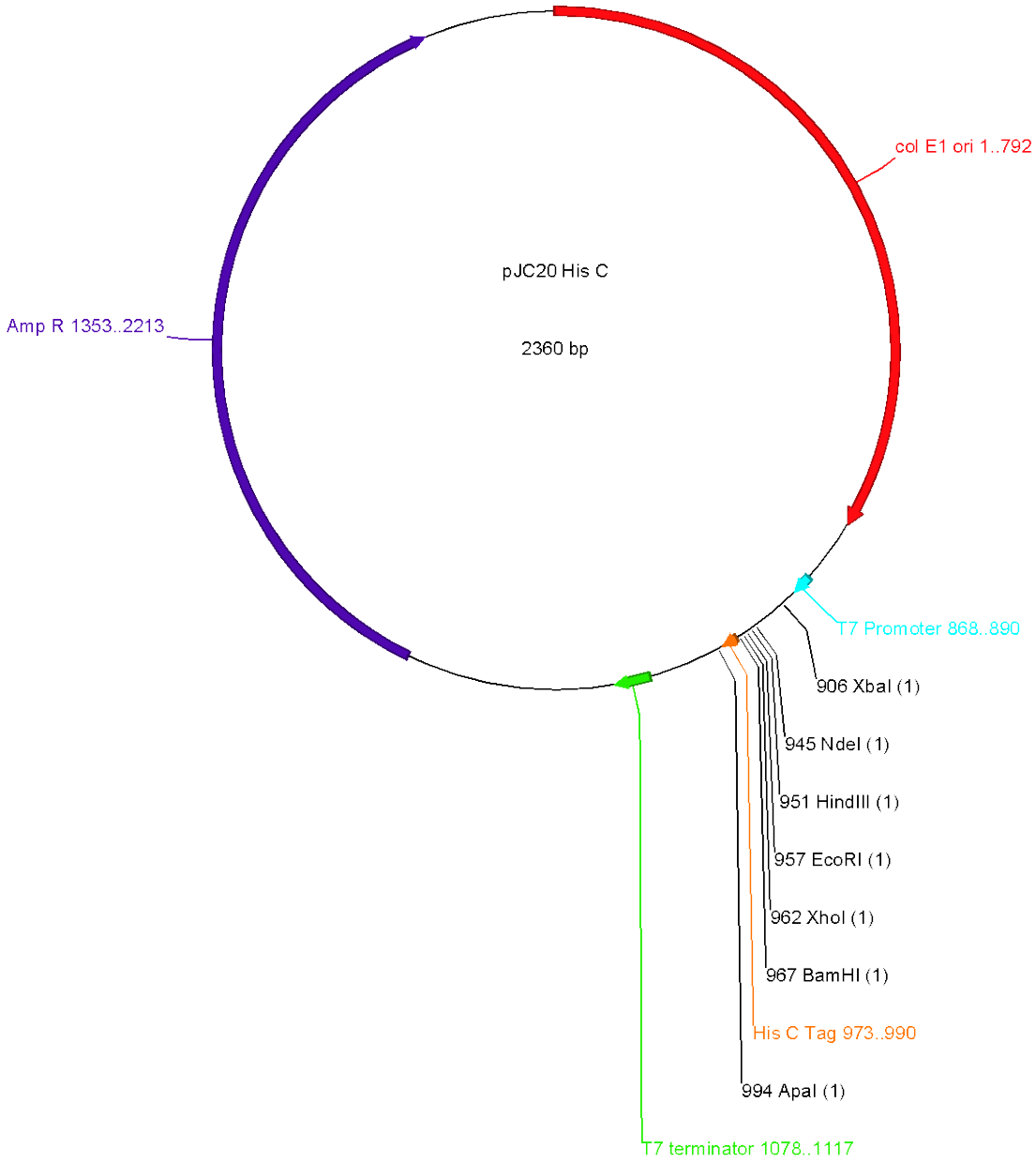
Appendix 9: TPI-GAPDH E165A Sequence

ATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCA
TATGTTCGTCGGTGGCAATTGGAAGTGCAACGGTTCTCTCTCCAAGGTCCAGGAGATCG
TGGCTACTCTCAACAACTCCAACCTCAACAACGACGCTGAGGTCGTGATCGCCCCTCCC
ACCGCCTATCTGAGAGACACCGTGGCTACCGTGCGCGCTGATGTGCAGGTGGCTGCTCA
GGACGTCTGGTCCCAGGGCAACGGTGCCTTCACCGGCGAGACCTCTGCTGAAATGCTGA
AGGATCTCAAGGTCGGTTGGGCGATCGTGGGCCACTCCGAGCGCCGTGGCAAGGGCGAG
AGCGATGCCGAGGTTGCGACCAAGGCTGCTTACGCTCAGAAGAACGGTCTGAAGGTGAT
CTGCTGCCTGGGCGAGTCCCTGAAGGAGCGCGAGGCTGGCCGCTTCGCCGAGGTGGTGA
CCCGCCAGCTCAAGGCCTACGCCGACGCCATCAAGAACTGGGATGACGCCGTGATCGCC
TACGCGCCCATCTGGGCCATTGGCACGGGCAAGACCGCCACCCCCGCTCAGGCGGAGGA
GGTGCACGCCGTACTGCGCAAGTGGCTGCGCGACAATGTGAGCGCCGCCGTGGCGGACA
AGGTGCGCATCATCTACGGAGGCTCCGTGAACGCGAAGAACTGCAACGAGCTGGGTAAG
CAGGCCGACATCGACGGCTTCCTGGTGGGCGGCGCTTCGCTGAAGCCCGAGTTCGTGAA
CATCATCAACGCCAACTCCACGGTGCAAAACGCGGGCCCCGTGAGCGTGGGTATCAACG
GCTTCGGCCGCATTGGTCGTCTGGTGTTCCGCGCCAGCCAGCCGAACCCGCTGGTGAAC
GTGGTGGCGATCAACGACCCGTTCATCACGCCCGACTACATGGAGTACATGATCATGCA
CGACTCCACGCACGGTCCGTTCCAGGGCACCGTGAAGGCTGAGAAGGACGCGATCATCG
TGAACGGCCGCCGCATCGTGGTGTCCAACGAGATGGACCCGAAGAAGATCCAGTGGGGC
GCTGCGGGTGCGGAGTACATCGTGGAGTCCACGGGCGTGTTCACGGCGAAGGACAAGGC
CGCGCAGCACCTGGAGGGCGGCGCGAAGAAGGTGGTGATTTCCGCGCCGTCGAAGGACG
CTCCGATGTTCGTGATGGGCGTGAACAACACGACGTACACGAAGAACCTGAAGGTGGTG
TCGAACGCGTCCTGCACGACGAACTGCCTGGCGCCGCTGGCGAAGATCGTGAATGACAA
GTTCGGTCTGAAAGAGGGTCTGATGACGACGGTGCACTCGGTGACCTCCACGCAAAAGG
TGCTGGACGGTCCTTCCAAGAAGGACTGGAGAGGCGGCCGCTCTGCCTGCTACAACATC
ATCCCGTCCTCCACGGGCGCCGCCAAGGCCGTGGGCAAGGTCATTCCCGAGCTGAACGG
CAAGCTGACGGGCATGTCCTTCCGCGTGCCGACGGAGGACGTCTCCGTGGTGGATCTGA
CCTGCACGCTGAAGAAGCCCGCCACCTACGAGCAGATCAAGGCTGCGGTGAAGGAGGCT
TCGGAGACGTACATGAAGGGCATTATGGGATACGTGGATTACGACGTGGTGTCTCGCG
ATCTGCTGACCTGCCCGTACTCTTCTGTCTTCGACGCGAAGGCAGGAATTGCCCTGAAC
GACACGTTCGTGAAGCTGGTTTCTTGGTACGACAACGAGTGGGGCTACTCCAACCGTAT
GGT

Appendix 10: TPI-GAPDH Sequence H96A E165A

ATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCA
TATGTTCGTCGGTGGCAATTGGAAGTGCAACGGTTCTCTCTCCAAGGTCCAGGAGATCG
TGGCTACTCTCAACAACTCCAACCTCAACAACGACGCTGAGGTCGTGATCGCCCCTCCC
ACCGCCTATCTGAGAGACACCGTGGCTACCGTGCGCGCTGATGTGCAGGTGGCTGCTCA
GGACGTCTGGTCCCAGGGCAACGGTGCCTTCACCGGCGAGACCTCTGCTGAAATGCTGA
AGGATCTCAAGGTCGGTTGGGCGATCGTGGGCGCCTCCGAGCGCCGTGGCAAGGGCGAG
AGCGATGCCGAGGTTGCGACCAAGGCTGCTTACGCTCAGAAGAACGGTCTGAAGGTGAT
CTGCTGCCTGGGCGAGTCCCTGAAGGAGCGCGAGGCTGGCCGCTTCGCCGAGGTGGTGA
CCCGCCAGCTCAAGGCCTACGCCGACGCCATCAAGAACTGGGATGACGCCGTGATCGCC
TACGCGCCCATCTGGGCCATTGGCACGGGCAAGACCGCCACCCCCGCTCAGGCGGAGGA
GGTGCACGCCGTACTGCGCAAGTGGCTGCGCGACAATGTGAGCGCCGCCGTGGCGGACA
AGGTGCGCATCATCTACGGAGGCTCCGTGAACGCGAAGAACTGCAACGAGCTGGGTAAG
CAGGCCGACATCGACGGCTTCCTGGTGGGCGGCGCTTCGCTGAAGCCCGAGTTCGTGAA
CATCATCAACGCCAACTCCACGGTGCAGAACGCGGGCCCCGTGAGCGTGGGTATCAACG
GCTTCGGCCGCATTGGTCGTCTGGTGTTCCGCGCCAGCCAGCCGAACCCGCTGGTGAAC
GTGGTGGCGATCAACGACCCGTTCATCACGCCCGACTACATGAAGTACATGATCATGCA
CGACTCCACGCACGGTCCGTTCCAGGGCACCGTGAAGGCTGAGAAGGACGCGATCATCG
TGAACGGCCGCCGCATCGTGGTGTCCAACGAGATGGACCCGAAGAAGATCCAGTGGGGC
GCTGCGGGTGCGGAGTACATCGTGGAGTCCACGGGCGTGTTCACGGCGAAGGACAAGGC
CGCGCAGCACCTGGAGGGCGGCGCGAAGAAGGTGGTGATTTCCGCGCCGTCGAAGGACG
CTCCGATGTTCGTGATGGGCGTGAACAACACGACGTACACGAAGGACCTGAAGGTGGTG
TCGAACGCGTCCTGCACGACGAACTGCCTGGCGCCGCTGGCGAAGATCGTGAATGACAA
GTTCGGTCTGAAAGAGGGTCTGATGACGACGGTGCACTCGGTGACCTCCACGCAAAAGG
TGCTGGACGGTCCTTCCAAGAAGGACTGGAGAGGCGGCCGCTCTGCCTGCTACAACATC
ATCCCGTCCTCCACGGGCGCCGCCAAGGCCGTGGGCAAGGTCATTCCCGAGCTGAACGG
CAAGCTGACGGGCATGTCCTTCCGCGTGCCGACGGAGGACGTCTCCGTGGTGGATCTGA
CCTGCACGCTGAAGAAGCCCGCCACCTACGAGCAGATCAAGGCTGCGGTGAAGGAGGCT
TCGGAGACGTACATGAAGGGCATTATGGGATACGTGGATTACGACGTGGTGTCTCGCG
ATCTGCTGACCTGCCCGTACTCTTCTGTCTTCGACGCGAAGGCAGGAATTGCCCTGAAC
GACACGTTCGTGAAGCTGGTTTCTTGGTACGACAACGAGTGGGGCTACTCCAACCGTAT
GGTCGACCTCATCCAGTACATGGCGAAGGTGGATCGCTCTTAA

Appendix 11: pJC20 His C Vector Map



Genes were inserted between NdeI and BamHI restriction sites.

Appendix 12: Prot Param data for TPI-GAPDH

Number of amino acids: 603

Molecular weight: 65012.0

Theoretical pI: 8.36

Amino acid composition:

Ala (A)	64	10.6%
Arg (R)	22	3.6%
Asn (N)	34	5.6%
Asp (D)	34	5.6%
Cys (C)	9	1.5%
Gln (Q)	16	2.7%
Glu (E)	29	4.8%
Gly (G)	51	8.5%
His (H)	13	2.2%
Ile (I)	32	5.3%
Leu (L)	34	5.6%
Lys (K)	45	7.5%
Met (M)	15	2.5%
Phe (F)	15	2.5%
Pro (P)	22	3.6%
Ser (S)	41	6.8%
Thr (T)	35	5.8%
Trp (W)	10	1.7%
Tyr (Y)	18	3.0%
Val (V)	64	10.6%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%

(B)	0	0.0%
(Z)	0	0.0%
(X)	0	0.0%

Total number of negatively charged residues (Asp + Glu): 63
Total number of positively charged residues (Arg + Lys): 67

Atomic composition:

Carbon	C	2869
Hydrogen	H	4558
Nitrogen	N	800
Oxygen	O	874
Sulfur	S	24

Formula: $C_{2869}H_{4558}N_{800}O_{874}S_{24}$
Total number of atoms: 9125

Extinction coefficients:

Extinction coefficients are in units of M^{-1} cm^{-1}, at 280 nm measured in water.

Ext. coefficient 82320
Abs 0.1% (=1 g/l) 1.266, assuming all pairs of Cys residues form cystines

Ext. coefficient 81820
Abs 0.1% (=1 g/l) 1.259, assuming all Cys residues are reduced

Estimated half-life:

The N-terminal of the sequence considered is M (Met).

The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).
>20 hours (yeast, in vivo).
>10 hours (Escherichia coli, in vivo).

Instability index:

The instability index (II) is computed to be 22.59
This classifies the protein as stable.

Aliphatic index: 84.08

Grand average of hydropathicity (GRAVY): -0.177

Appendix 13: Prot Param data for TPI

Number of amino acids: 267

Molecular weight: 28556.1

Theoretical pI: 6.76

Amino acid composition:

Ala (A)	36	13.5%
Arg (R)	11	4.1%
Asn (N)	18	6.7%
Asp (D)	13	4.9%
Cys (C)	4	1.5%
Gln (Q)	9	3.4%
Glu (E)	16	6.0%
Gly (G)	24	9.0%
His (H)	9	3.4%
Ile (I)	13	4.9%
Leu (L)	16	6.0%
Lys (K)	17	6.4%
Met (M)	3	1.1%
Phe (F)	5	1.9%
Pro (P)	6	2.2%
Ser (S)	17	6.4%
Thr (T)	12	4.5%
Trp (W)	6	2.2%
Tyr (Y)	5	1.9%
Val (V)	27	10.1%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%

(B)	0	0.0%
(Z)	0	0.0%
(X)	0	0.0%

Total number of negatively charged residues (Asp + Glu): 29
Total number of positively charged residues (Arg + Lys): 28

Atomic composition:

Carbon	C	1248
Hydrogen	H	1980
Nitrogen	N	368
Oxygen	O	387
Sulfur	S	7

Formula: $C_{1248}H_{1980}N_{368}O_{387}S_7$
Total number of atoms: 3990

Extinction coefficients:

Extinction coefficients are in units of M^{-1} cm^{-1}, at 280 nm measured in water.

Ext. coefficient 40700
Abs 0.1% (=1 g/l) 1.425, assuming all pairs of Cys residues form cystines

Ext. coefficient 40450
Abs 0.1% (=1 g/l) 1.417, assuming all Cys residues are reduced

Estimated half-life:

The N-terminal of the sequence considered is M (Met).

The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).
>20 hours (yeast, in vivo).
>10 hours (Escherichia coli, in vivo).

Instability index:

The instability index (II) is computed to be 16.30
This classifies the protein as stable.

Aliphatic index: 85.17

Grand average of hydropathicity (GRAVY): -0.249

Appendix 14: Prot Param data for GAPDH

Number of amino acids: 356

Molecular weight: 38711.4

Theoretical pI: 8.88

Amino acid composition:

Ala (A)	28	7.9%
Arg (R)	12	3.4%
Asn (N)	16	4.5%
Asp (D)	21	5.9%
Cys (C)	5	1.4%
Gln (Q)	7	2.0%
Glu (E)	13	3.7%
Gly (G)	29	8.1%
His (H)	11	3.1%
Ile (I)	19	5.3%
Leu (L)	19	5.3%
Lys (K)	28	7.9%
Met (M)	14	3.9%
Phe (F)	10	2.8%
Pro (P)	17	4.8%
Ser (S)	29	8.1%
Thr (T)	23	6.5%
Trp (W)	4	1.1%
Tyr (Y)	13	3.7%
Val (V)	38	10.7%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%

(B)	0	0.0%
(Z)	0	0.0%
(X)	0	0.0%

Total number of negatively charged residues (Asp + Glu): 34
Total number of positively charged residues (Arg + Lys): 40

Atomic composition:

Carbon	C	1714
Hydrogen	H	2717
Nitrogen	N	469
Oxygen	O	513
Sulfur	S	19

Formula: $C_{1714}H_{2717}N_{469}O_{513}S_{19}$
Total number of atoms: 5432

Extinction coefficients:

Extinction coefficients are in units of M^{-1} cm^{-1}, at 280 nm measured in water.

Ext. coefficient 41620
Abs 0.1% (=1 g/l) 1.075, assuming all pairs of Cys residues form cystines

Ext. coefficient 41370
Abs 0.1% (=1 g/l) 1.069, assuming all Cys residues are reduced

Estimated half-life:

The N-terminal of the sequence considered is M (Met).

The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).
>20 hours (yeast, in vivo).
>10 hours (Escherichia coli, in vivo).

Instability index:

The instability index (II) is computed to be 28.99
This classifies the protein as stable.

Aliphatic index: 80.45

Grand average of hydropathicity (GRAVY): -0.173

Appendix 15: Average initial rates of reaction with GAP as substrate

Concentration mM	Average initial rate of enzyme (mM mM^{-1} s^{-1})						
	wtTPI-GAPDH	GAPDH His N	GAPDH His C	TPI-GAPDH C404A	TPI-GAPDH H96A	TPI-GAPDH E165A	TPI-GAPDH H96A E165A
0.01	0.0185	0.0044	0.0110	0.0041	0.0863	0.0635	0.0854
0.03	0.0389	0.0064	0.0170	0.0032	0.1690	0.1274	0.1871
0.06	0.0988	0.0216	0.0292	0.0051	0.3548	0.3008	0.3955
0.13	0.1871	0.0238	0.0563	0.0064	0.7280	0.5977	0.7820
0.25	0.3522	0.0396	0.0928	0.0188	1.1269	0.9608	1.3805
0.5	0.5822	0.0500	0.1791	0.0207	1.6832	1.6936	2.0446

Appendix 16: Average initial rates of reaction with DHAP as substrate

Concentration mM	Average initial rate of enzyme (mM mM^{-1} s^{-1})					
	wtTPI-GAPDH	GAPDH His C	TPI-GAPDH C404A	TPI-GAPDH H96A	TPI-GAPDH E165A	TPI-GAPDH H96A E165A
0.01	0.0264	No other concentrations assayed as rate so low	0.0022	0.0021	0.0010	0.0013
0.03	0.0322		0.0027	0.0031	0.0021	0.0019
0.06	0.1121		0.0047	0.0052	0.0036	0.0036
0.13	0.1813		0.0068	0.0065	0.0062	0.0052
0.25	0.3453	0.0002	0.0108	0.0125	0.0084	0.0066
0.5	0.5042		0.0151	0.0181	0.0135	0.0121

Appendix 17: Bacterial strains used in study

Strain	Genotype	Source
XL1-Blue	wtTPI-GAPDH His-6 N PET14b	M vdGiezen
XL1-Blue	GAPDH His-6 N PET14b	M vdGiezen
XL1-Blue	TPI His-6 N PET14b	M vdGiezen
XL1-Blue	TPI-GAPDH His-6 N PET14b C404A	This project
XL1-Blue	TPI-GAPDH His-6 N pJC20 H96A	This project
XL1-Blue	TPI-GAPDH His-6 N pJC20 E165A	This project
XL1-Blue	TPI-GAPDH His-6 N pJC20 H96A E165A	This project
XL1-Blue	GAPDH His-6 C pJC20	This project
BL21 DE3 plysS	wtTPI-GAPDH His-6 N PET14b	This project
BL21 DE3 plysS	GAPDH His-6 N PET14b	This project
BL21 DE3 plysS	TPI His-6 N PET14b	This project
BL21 DE3 plysS	TPI-GAPDH His-6 N PET14b C404A	This project
BL21 DE3 plysS	TPI-GAPDH His-6 N pJC20 H96A	This project
BL21 DE3 plysS	TPI-GAPDH His-6 N pJC20 E165A	This project
BL21 DE3 plysS	TPI-GAPDH His-6 N pJC20 H96A E165A	This project
BL21 DE3 plysS	GAPDH His-6 C pJC20	This project